

GTGTTTATGATGATAAAAAGGACAGCAAGTAATGATGATACTAAGCAAA  
TCTTACTCTCTACGTAAATATAAATTTGGTTTAGCATCAGTAA  
TTTTAGGGTCATTTCATAATGGTCACAAGTCCTGTTTTTGCGGA  
TCAAACCTACATCGGTTCAAGTTAATAATCAGACAGGCACTAG  
TGTGGATGCTAATAATTCTTCCAATGAGACAAGTGCCTCAAGT  
GTGATTACTTCCAATAATGATAGTGTTC AAGCGTCTGATAAAG  
TTGTAATAAGTCAAAATACGGCAACAAAGGACATTACTACTC  
CTTTAGTAGAGACAAAGCCAATGGTGGA AAAACATTACCTG  
AACAAGGGAATTATGTTTATAGCAAAGAAACCGAGGTGAAAA  
ATACACCTTCAAAATCAGCCCCAGTAGCTTTCTATGCAAAGAA  
AGGTGATAAAGTTTTCTATGACCAAGTATTTAATAAAGATAAT  
GTGAAATGGATTTTCATATAAGTCTTTTGGTGCGGTACGTCGAT  
ACGCAGCTATTGAGTCACTAGATCCATCAGGAGGTTTCAGAGA  
CTAAAGCACCTACTCCTGTAAACAAATTCAGGAAGCAATAATC  
AAGAGAAAAATAGCAACGCAAGGAAATTATACATTTTCACATA  
AAGTAGAAGTAAAAAATGAAGCTAAGGTACGAGGTCCAACCTC  
AATTTACATTGGACAAGGAGACAGAATTTTTTACGACCAAAA  
TACTAACTATTGGAAGGAATCAGTGGTTATCTTATAAATCATT  
CAATGGTGTTTCGTCGTTTTGTTTTGCTAGGTAAAGCATCTTCA  
GTAGAAAAAACTGAAGATAAAGAAAAAGTGTCTCCTCAACCA  
CAAGCCCGTATTACTAAAACCTGGTAGACTGACTATTTCTAACG  
AAACAACCTACAGGTTTTGATATTTTAATTACGAATATTAAGA  
TGATAACGGTATCGCTGCTGTAAAGGTACCGGTTTGGACTGAA  
CAAGGAGGGCAAGATGATATTAATGGTATACAGCTGTAACCT  
ACTGGGGATGGCAACTACAAAGTAGCTGTATCATTTGCTGAC  
CATAAGAATGAGAAGGGTCTTTATAATATTCATTTATACTACC  
AAGAAGCTAGTGGGACACTTGTAGGTGTAACAGGAACTAAAG  
TGACAGTAGCTGGAACCTAATTCCTCTCAAGAACCTATTGAAAA  
TGGTTTACCAAAGACTGGTGTTTATAATATTATCGGAAGTACT  
GAAGTAAAAAATGAAGCTAAAATATCAAGTCAGACCCAATTT  
ACTTTAGAAAAAGGTGACAAAATAAATTATGATCAAGTATTG  
ACAGCAGATGGTTACCAAGTGGATTTCTTACAAATCTTATAGTG  
GTGTTTCGTCGCTATATTTCCTGTGAAAAAGCTAACTACAAGTAG  
TGAAAAAAGCGAAAAGATGAGGCGACTAAACCGACTAGTTATCC  
CAACTTACCTAAAACAGGTACCTATACATTTACTAAAACCTGTA  
GATGTGAAAAGTCAACCTAAAGTATCAAGTCCAGTGGAATTT  
AATTTTCAAAAGGGTGAAAAAATACATTATGATCAAGTGTTA  
GTAGTAGATGGTCATCAGTGGATTTTCATACAAGAGTTATTCCG  
GTATTCGTCGCTATATTGAAATTTAA

2 / 110

MFMMKKGQVNDTKQSYSLRKYKFGLASVILGSFIMVTSPVFADQTTSVQVNN  
 QTGTSVDANNSSNETSASSVITSNNDVQASDKVVNSQNTATKDITPLVETK  
 PMVEKTLPEQGNVYVYSKETEVKNTPSKSAPVAFYAKKGDKVFYDQVFNKDN  
 VKWISYKSFGGVRRYAAIESLDPSGGSETKAPTPVTNSGSNNQEKIATQGNYT  
 FSHKVEVKNEAKVASPTQFTLDKGDRIFYDQILTIEGNQWLSYKSFNGVRRFV  
 LLGKASSVEKTEDKEKVSPQPQARITKTGRLTISNETTTGFDILITNIKDDNGIA  
 AVKVPVWTEQGGQDDIKWYTAVTTGDGNYKVAVSFADHKNEKGLYNIHLY  
 YQEASGTLVGVGTGKVTVAGTNSSQEPINGLPKTGVYNIIGSTEVKNEAKISS  
 QTQFTLEKGDKNYDQVLTADGYQWISYKSYSGVRRYIPVKKLTSSEKAKDE  
 ATKPTSYPNLPKTGTYTFTKTVDVKSQPKVSSPVEFNFQKGEKIHYDQVLVVD  
 GHQWISYKSYSGIRRYIEI\*

Sequence description

A) Length: 1642 bp - 547 aa (full length gene)

B) Sequence Characteristics:

Potential leader peptide sequence

Orf is preceded by a potential Shine-

Dalgarno sequence.

ID-66

Clone 3-5

ATGATATTGAGACGTCGAACTATTGTTTTATGGCAACTGGGTATCGCCATT  
 TCTCTCATTCTTAGTATTCTAGCCTTAAATCTTTATTTCCATAGTACTCCCTT  
 GCAAACCAATGCAGCTTTACGGAACCTTGCTCCTTCATTAAACCATCTTTTT  
 GGGACAGATGGTTTAGGTAGGGATATGTTTGTGAGAACGATTAAAGGACT  
 TTATTTCTCTCTACAAGTCGGCTTATTAGGTGCCCTTATGGGGGTCATTCTG  
 GCGACAGTTTTTGGAGTGCTTGCAGGTTTAGGAAATAGCATTATTGATAAA  
 ATAATAGCATGGTTAGTTGATTGTTTTATTGGTATGCCTCATTTGATTTTGA  
 TGATTCTCATTCTTTTTGTTGTTGGGAAAGGTGCTCAAGGGGTCATCATTGC  
 AACGGCTGTTACACATTGGCCTTCTTTAGCAAGGCTTATCCGCAATGAAGT  
 CTATCATCTAAAGAATAAAGAATTTGTCCAACCTTTCTAAAAGTATGGGAAA  
 AACGCCTTATTATATTGTGAGGCATCATATCCTGCCTTTGATTGCTTCTCAA  
 ATTTTCATTGGTTTTATCCTCTTATTTCCACATGTCATCCTACATGAAGCAT  
 CAATGACTTTCTTAGGATTTGGGCTCTCTGCCGAACAACCTTCGGTTGGTA  
 TCATTCTGTCAGAGGCAGCTAAGCATATCTCTCTTGGAATTTGGTGGTTGG  
 TTATCTTTCCAGGACTTTATCTTATTTTGGTTGTCAATGCATTTGATACTAT  
 CGGAGAATCTTTAAAGAACTCTTTTACCCTCAAACCTGATCATTTTTAG

FIG. 1CONT'D

3 / 110

MILRRRTIVLWQLGIAISLILSILALNLYFHSTPLQTNAALRNLAPSLNHLFGTD  
 GLGRDMFVRTIKGLYFSLQVLLGALMGVILATVFGVLAGLGNHDKIIAWL  
 VDLFIGMPHLIFMILISFVVVGKGAQGVIIATAVTHWPSLARLIRNEVYHLKNKE  
 FVQLSKSMGKTPYYIVRHHILPLIASQIFIGFILLFPHVILHEASMTFLGFGLSAE  
 QPSVGILSEAAKHISLGNWWLVIFPGLYLILVNAFDTIGESLKKLFYPQTDHF  
 \*

#### Sequence description

- A) Length: 822 bp - 274 aa (full length gene)
- B) Sequence Characteristics:  
 Potential leader peptide sequence  
 Orf is preceded by a potential Shine-  
 Dalgarno sequence.

ID-78

Clone 3-5b

ATGACAGAAACATTATTAAGCATTAAAGACCTCTCCATCACCTTCACTCAA  
 TACGGAAGATTTTTAAAACCATTTCAATCAACACCGATACAAGCGCTGA  
 ATTTAGAAATTAAAAAAGGTGAGTTATTAGCTATTATAGGTGCTAGTGGTT  
 CGGGGAAGAGTTTATTAGCACATGCTATTATGGATATTCTTCCTAAAAATG  
 CATCTGTAACAGGAGATATGATTTATCGTGGTCAATCACTAAATTCTAAAC  
 GCATTAAACAGTTGCGAGGAAAAGATATTACGTTGATTCCACAATCAGTTA  
 ATTATTTAGATCCATCTATGAAAGTCAAACATCAGGTGCGCTTAGGTATCT  
 CAGAAAATTCAAAGGCTACTCAAGAAGGATTGTTTCAACAGTTTGGTTTAA  
 AAGAAAGTGATGGTGACTTGGATCCTTTCCAACCTTTCTGGCGGAATGCTCC  
 GACGTGTTTTGTTTACAACGTGTATTAGTGATAAGGTTTCTTTGATTATTGC  
 GGATGAGCCCACCCCTGGATTACATCCAGATGCTCTGCAAATGGTTTTAGA  
 CCAACTACGCTCCTTTGCAGATAAAGGAATAAGCGTTATATTTATCACTCA  
 TGATATTGTAGCAGCTAGTCAAATTGCTGATCGTATTACTATTTTAAAGA  
 GGGAAAAGCTATTGAAACAGCTCCAGCTAGTTTCTTTAGCGGAAATGGAG  
 AGCAGTTACAAACAGAATTTGCTAGAAAGTTTATGGCGCTCTCTCCACAGC  
 AAGAATTTTTGAAAGGAGTTACTCATGACCTTAGAGGCTAA

MTETLLSIKDLSITFTQYGRFLKPFQSTPIQALNLEIKKGELLAHIGASGSGKSL  
 AHAIMDILPKNASVTGDMYRQSLNSKRIKQLRGKDITLIPQSVNYLDPSMK  
 VKHQVRLGISENSKATQEGLFQQFGLKESDGDLDPFQLSGGMLRRVLFTTCIS  
 DKVSLIIADEPTPGLHPDALQMVLDQLRSFADKGISVIFITHDIVAASQIADRITI  
 FKEGKAJETAPASFFSGNGEQLQTEFARSLWRSLPQQEFLKGVTHDLRG\*

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

FIG. 1 CONT'D

TTGAAAAATTTAAATCGTTATGTAGTTGCGGTTTCTGGAGTCGTTTTACATT  
 TAATGCTAGGATCAACTTATGCTTGGAGTGTGTTTCGTAACCCAATTATCT  
 CAGAGACTGGTTGGGATATTCATCAGTTTCATTCGCTTTTAGTTTGGCTAT  
 TTTTGTCTAGGAATGTCTGCAGCTTTTATGGGACACTTAGTAGAGCGTTTT  
 GGTCTAGGATAATGGGAATGATTTCTGCTATTTTATATGGAGCAGGGAAT  
 GTGTTAACAGGCTTAGCCATTGAAACTCAGCAGTTATGGTTACTGTATGTT  
 GCATACGGTATTTAGGAGGAATCGGACTTGGTTCAGGTTATATTACTCCA  
 GTATCGACTATTATTAATGGTTTCCTGATAGGAGGGGACTAGCAACAGG  
 ATTCGCTATTATGGGATTTGGCTTTGCTTCTTTAGTAACAAGTCCGCTTGCA  
 CAATCCTTACTGATTAGGATTGGTGTGGGTAACGTTTTATATTTTGGGA  
 TTAGTATATTTTGTGTCATGATGATTGCCTCACAATTTATTAACAACCAC  
 CTCAGGAAAAAATAACTATTTGACTCACGATGGTAAAAAGAATGCTATG  
 AATTCACAAATTATCACTGGATTAAGCAACGTCGCTATAAAATCAAA  
 AACCTTTTACATCATTTGGTTGACCTTGTATTATTAATATTTCTGTGTGGCTTA  
 GGTTTAATATCAGCAGCTTCACCAATGGCACAAGATTTAGCAGGCTATTCC  
 GCAGAATCTGCAGCCTTATTAGTAGGGGTACTAGGGATATTTAACGGTTTT  
 GGACGTCTGTTATGGGCAAGTCTCTCTGACTACATTGGACGCCCCGTTGACC  
 TTTATAATATTATTTATTGTGAACCTTATTATGACTTCTAGTTTATTTTGTG  
 ATTCAATGCTATTGTATTTGCAATAGCGATGTCTATTTTAATGACTTGTTAT  
 GGTGCAGGTTTTTCTTATTACCTGCTTATCTAAGTGATATTTTGGAAACA  
 AGGAATTAGCTACTTTACATGGTTATAGTTTAAACAGCATGGGCAATAGCAG  
 GTCTGTTTGGGCCCCCTATTGTTATCAAAGACATATTCATGGGGAAATTCCT  
 ATCAATTGACATTAATGGTTTTTGGTTTTTTATTCTTATTCGATTATTGTTA  
 TCTCTATATTTAAGAAAATTAACAACCTAAAGTTGTGTAG

LKNLNRYVVAVSGVVLHMLLGSTYAWSVFRNPIISETGWDISSVSFAFLAIFC  
 LGMSAAFMGHLVERFGPRIMGMISAILYGAGNVLTGLAIETQQLWLLYVAYG  
 ILGGIGLGSYITPVSTIHKWFPDRRGLATGFAIMGFGFASLVTSPLAQSLIRIG  
 VGKTFYILGLVYFFVMMIASQFIKQPPQEKITILTHDGKKNAMNSQIITGLKAN  
 VAIKSKTFYIWLTLFINISCGLGLISAASPMQDLAAGYSAESAALLVGVLGIFN  
 GFGRLLWASLSDYIGRPLTFIILFIVNFIMTSSLFLSFNAIVFAIAMSILMTCYGA  
 GFSLLPAYLSDIFGTELATLHGYSLTAWAIALGLFGPLLLSKTYSWGNSYQLTL  
 MVFGFLFLFGLLLSLYLRKLTTKVV\*

Sequence description:

- A] Length 1221 bp - 407 a.a (full length gene).
- B] TTG start codon with Shine-Dalgarno sequence upstream. Obvious signal peptide, with hydropathy plot exhibiting many possible membrane spanning regions, indicating protein to be transmembrane.

FIG. 1 CONT'D

7 / 110

ID-82

Clone 48

ATGGCAGATAAAAAACAGAACATTTAAACTTGTAGGTGCAGGATCTTCTAG  
CACACAAGAAAAAATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTG  
GCGTCGCTTGAAAAAAAACAAATTAGCAGTAGTTTCACTCTATTTATTAGC  
TCTTTTACTTACTTTTTTCGTTAGCCTCAAATTTATTTGTAACTCAGAAGGAT  
GCTAATGGGTTTGATTGAAAAAAGTAACGACATATCGCAACTTACCACCT  
AAATTGAGTTCAAACCTTCCTTTTTTGAATGGTAGCATTAAATCCATCA

MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLA  
LLLTFSLASNLFTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSINPS

Sequence description:

A) Current length is 303 bp - 101 aa

B) No obvious signal peptide but Shine

Dalgarno sequence upstream of the ATG start

codon. Not identified directly using the LEEP system but was found

directly downstream of ID-34 described in WO 00/06736.

ID-83

Clone 98

ATGAAAATAGTAGTACCAGTAATGCCTCGCAGTCTTGAAGAGGCTCAAGA  
AATAGATTTATCAAAATTTGATAGTGTTGATATTATTGAATGGCGAGCTGA  
TGCTTACCAAAGGATGACATTATTAATGTAGCTCCAGCTATTTTTGAGAA  
ATTCGCAGGTCATGAAATTATTTTTACTTTTCGTACAACGCGTGAAGGTGG  
TAATATTGTCTTATCTGATGCTGAGTATGTTGAGTTAATCCAGAAAATTAA  
TTCTATCTACAATCCAGATTATATTGATTTTGAAGTATTTTTCACATAAAGAA  
GTTTTTCAAGAAATGCTAGAATTTCCAAATTTAGTCCTGTCTTATCACAATT  
TTCAAGAGACACCGGAGAATATTATGGAGATATTTTCAGAAATTAACAGCC  
CTAGCACCACGAGTTGTGAAAATCGCAGTAATGCCAAAGAATGAACAAGA

FIG. 1<sup>CONT'D</sup>

TGTCTTAGACGTTATGAATTACACTCGCGGTTTCAAGACTATTAATCCTGA  
TCAAGTTTATGCGACGGTATCTATGAGTAAAATTGGACGTATTTCTCGTTTT  
GCTGGTGATGTAAGTGGATCTAGTTGGACATTTGCATATTTAGATTCATCT  
ATCGCACCCGGACAAATTACTATTTTCAGAGATGAAGCGTGTCAAAGCATT  
GCTTGACGCTGACTGA

MKIVVPVMPRSLEEAQEIDLSKFDSVDIIEWRADALPKDDIINVAPAIFEFKAG  
HEIIFTFRTTREGGNIVLSDAEYVELIQKINSIYNPDYIDFEYFSHKEVFQEMLEF  
PNLVLSYHNFQETPENIMEIFSELTALAPRVVKIAVMPKNEQDVLDVMNYTRG  
FKTINPDQVYATVSMKIGRISRFAAGDVTGSSWTFAYLDSSIAPGQITISEMKRV  
KALLDAD\*

Sequence description:

- A] Length 678 bp, 225 aa (full length gene)
- B] No obvious signal peptide, but there is a Shine Dalgarno immediately upstream of ORF.

ID-84

Clone RS-52

ATGAAAGACTTATTTGCAACAACAGAAGCATCATCAAGGAAACAGGAACA  
AGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAACAAATGGTA  
ATCAAATAAAAAAAAAATTGAGTTTATCGACTTTCAAAAAAATGAGATGACA  
GGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCGATTAGTTTT  
TCTGAAGATAGAATTGGTGGTAACTTAGAGCATTAGGATATCAACCGAA  
TGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCAAATGTTAATGA  
TATTGAAGTGATTTATATGAAGAAAGAATAG

MKDLFATTEASSRKQEQDRIVNYIKQHVELTNGNQIKKIEFIDFQKNEMTGTW  
GISTKINEQFSISFSEDRI GGKLRALGYQPNEIGFSKDINSNNQNVNDIEVIYMK  
KE\*

Sequence description:

- A] length: 333 bp - 111 aa (partial sequence)
- B] No obvious Shine Dalgarno sequence upstream of the ATG start codon, and no obvious signal peptide within the protein.

FIG. 1<sub>CONT'D</sub>

9 / 110

ID-85

Clone RS-53

ATGAAAAACGTATATGGTATTTGATAATAATAATCACAGTAATTTTAGGA  
GGACTAGCCATGAAAACTTATTTGCAACAACAGAAGCATCATCAAGGAA  
ACAGGAACAAGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAA  
CAAATGGTAATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAT  
GAGATGACAGGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCG  
ATTAGTTTTTCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATAT  
CAACCGAATGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCA

MKKRIWYLIHITVILGGLAMKNLFATTEASSRKQEQDRIVNYIKQHVELTNGN  
QIKKIEFIDFQKNEMTGTWGISTKINEQFSISFSEDRIGGKLRLALGYQPNEIGFSK  
DINSNNQ

Sequence description:

- A) Length: 351 bp - 117 aa (Partial sequence)
- B) Obvious signal peptide and Shine Dalgarno sequence upstream of the ATG start codon.

ID-86

Clone ID-74

ATGTCAAATCAATATGATTATATCGTTATTGGTGGAGGTAGT  
GCAGGCAGTGGTACCGCTAATAGGGCAGCCATGTATGGAGC  
AAAAGTCCTGTTAATTGAAGGTGGACAAGTAGGTGGAACCTG  
TGTTAACTTAGGTTGTGTACCTAAGAAAATCATGTGGTATGG  
TGCACAAGTTTCTGAGACACTCCATAAGTATAGTTTCAGGTTA  
TGGTTTTGAAGCCAATAATCTTAGTTTTGATTTTACTACTCTA  
AAAGCTAATCGCGATGCTTACGTGCAGCGGTCTAGACAGTCG  
TATGCCGCTAATTTTGAGCGTAATGGGGTCGAAAAGATTGAT  
GGATTTGCTCGTTTTATTGATAACCATACTATTGAAGTGAATG  
GTCAGCAATATAAAGCTCCTCACATTACTATTGCAACAGGTG

FIG. 1<sub>CONT'D</sub>

10 / 110

GACACCCTCTTTACCCTGATATTATTGGAAGTGAAGTGGTG  
 AGACTTCTGATGATTTTTTTGGATGGGAGACCTTACCAAATTC  
 TATATTGATTGTTGGGGCGGGCTATATCGCGGCAGAACTTGC  
 TGGAGTGGTTAATGAATTAGGCGTTGAAACCCATCTTGCATT  
 TAGAAAAGACCATATTCTACGCGGATTTGATGACATGGTAAC  
 AAGTGAGGTTATGGCTGAAATGGAGAAATCAGGTATCTCTTT  
 ACATGCTAACCATGTACCTAAATCTCTTAAACGCGATGAAGG  
 TGGCAAGTTGATTTTTGAAGCTGAAAATGGGAAAACGCTTGT  
 CGTTGATCGTGTAATATGGGCTATCGGCCGTGGACCAAATGT  
 AGACATGGGACTTGAAAATACCGATATTGTTTTAAATGATAA  
 AGATTATATCAAAACAGATGAATTTGAGAATACTTCTGTAGA  
 TGGCGTGTATGCTATTGGAGATGTTAATGGGAAAATTGCCTT  
 GACACCGGTAGCAATTGCAGCAGGTCGTCGCTTATCAGAAAG  
 ACTTTTAAATCATAAAGATAACGAAAAATTAGATTACCATAA  
 TGTACCTTCAGTTATTTTTACTCACCTGTAATTGGGACGGTA  
 GGACTTTCAGAAGCAGCAGCTATCGAGCAATTTGGAAAAGAT  
 AATATCAAAGTCTATACATCAACTTTTACCTCTATGTATACGG  
 CTGTTACCAGTAATCGCCAAGCAGTTAAGATGAAGCTCATAA  
 CCCTAGGAAAAGAGGAAAAAGTTATTGGGCTTCATGGTGTTG  
 GTTATGGTATTGATGAAATGATTCAAGGTTTTTCAGTTGCTAT  
 CAAAATGGGGGCTACTAAAGCAGACTTTGATGATACTGTTGC  
 TATTCACCCAACCTGGATCTGAGGAATTTGTTACAATGCGCTA  
 A

MSNQYDYIVIGGGSAGSGTANRAAMYGAKVLLIEGGQVGGTC  
 VNLGCVPKKIMWYGAQVSETLHKYSSGYGFEANNLSFDFTTLK  
 ANRDAYVQRSRQSYAANFERNGVEKIDGFARFIDNHTIEVNGQ  
 QYKAPHITATGGHPLYPDIIGSELGETSDDFFGWETLPNSILIVG  
 AGYIAAELAGVVNELGVETHLAFRKDHILRGFDDMTSEVMAE  
 MEKSGISLHANHVPKSLKRDEGGKLIFEANGKTLVVDRIWAI  
 GRGPNVDMGLENTDIVLNDKDYIKTDEFENTSVDGVYAIGDVN  
 GKIALTPVAIAAGRRLSERLFNHKDNEKLDYHNVPSVIFTHPVIG  
 TVGLSEAAAIEQFGKDNKIVYTSTFTSMYTAVTNRQAVKMKLI  
 TLGKEEKVIGLHGVGYGIDEMIQGFSVAIKMGATKADFDDTVAI  
 HPTGSEEFVTMR\*

ID-87

FIG. 1 CONT'D

FIG. 1 CONT'D

12 / 110

TTAAAAATAATCGAAGGTTTAGATTACTCTACATTACGTCAGACTACTCAA  
 ACTCAGTTTAGAGAAATGACTACTGCAGGTAAAGCGTTGTTATCAAAATCT  
 AACCTCCGACAAGGAGAAAAACAAAATTCTTCAAGAAGCACAATTTTT  
 CCTTGGTCGCGTTGATTTGGATAAAGCCATAGCTAAAGCTGAGAAGGCTTT  
 AGTGACCAAGAAGGCAACAAAGAATGGTCATTTGCTTGAGAGGAGTATTA  
 ACAAAGCGGTATTAGCTTATAATAATAGTGCTATTAAGGCTAATGTTA  
 AGCGCTTGAAAAAGAGTTAGACTTGCTGACAGATTTAGTCGAGGGAAAA  
 GGACCATTAGCGCAAGCTACAATGGTACAAGGAGTTTATTTATTAAGAC  
 GCCTTTACCATTGCCAGAATATTATATCGGATTGAACGTTTATTTTGACAA  
 GTCTGGAAAATTGATTTATGCACCTTGATATGAGTGATACTATTGGCGAGGG  
 AAAAAAGATGCATATGGTAATCCTATATTAATGTTGACGAGGATAATG  
 AAGGTTATCATACCTTGGCAGTTGCCACTTTAGCTGATTATGAAGGTCTTT  
 ATATTAAGATATTTTAAATAGTTCCTTGATAAGATTAAAGCAATACGCC  
 AGATTCTTTGGCAAAATATCATAGATTAGGAATTTCCAAGCTATCCGAA  
 ATGCAGCGGCAGAAGCAGACCGATTGCTTCCTAAGACACCTAAGGGGTAC  
 CTAATGAAGTCCCAAATTATCGTAAAAAACAAATGGAGAAAAATTTAAA  
 ACCAGTTGATTATAAAACGCCGATTTTAATAAGGCTTTACCTAATGAAAA  
 GGTAGACGGTGATAGAGCGGCTAAAGGTCATAATATAAATGCGGAGACTA  
 ATAATTCTGTAGCTGTAAACACCAATAAGGTCCGAGCAGCAATTACATAAGT  
 CACAGTCTGATGTAAATTTACCTCAAACAAGTTCTAAAAATAATTTTATAT  
 ACGAGATTCTAGGATACGTTAGTTTATGTTTGCTTTTCCTAGTAAGTCTGG  
 GAAAAAGGAAAACGAGCAAGAAAAATAA

MTKKHLKTLALATTVSVVTYSQEVYGLEREESVKQEQTQSASEDDWFEEDN  
 ERKTNVSKENSTVDETVSDLFSDGNSNNSSSKTESVVS DPKQVPKAKPEVTQE  
 ASNSSNDASKVEVPKQDTASKKETLETSTWEAKDFVTRGDTLVGFSKSGINKL  
 SQTSHLVLP SHAADGTQLTQVASFAFTPDKKTAAEYTSRLGENGKPSRLDIDQ  
 KEIIDEGEIFNAYQLTKLTIPNGYK SIGQDAFVDNK NIAEVNLPESLETISDYAF  
 AHMSLKQVKLPDNLKVIGELAFFDNQIGGKLYLPRHLIKLAERAFKSNRIQTV  
 EFLGSKLK VIGEASFQDNNLRNVMLPDGLEKJESAEFTGNPGDEHYNNQVVL R  
 TRTGQNP HQ LATENTYVNPDKSLWRATPDMDYTKWLEEDFTYQKNSVTGFS  
 NKGLQKVRRNKNLEIPKQHNGITITEIGDNAFRNVDFQSKTLRKYDLEEIKLPS  
 TIRKIGAFAFQSNNLKSFEASEDLEEIKEGAFMNNRIGTLDLKD KLIKIGDAAFH  
 INHIYAI VLPESVQEIGRSAFRQNGALHLMFIGNKVKTIGEMAFLSNKLESVNL  
 SEQKQLKTIEVQAFSDNALSEVVLPPNLQTIREEAFKR NHLKEVKGSSTLSQITF  
 NAFDQNDGDKRFGKKVVVRTHNNSHMLADGERFIIDPKLSSTMVDLEKVL  
 KIEGLDYSTLRQTTQTQFREMTTAGKALLSKSNLRQGEKQKFLQEAQFFLGR  
 VDLDKAIKAEKALVTKKATKNHLLERSINKAVLAYNNSAIKKANVKRLEK  
 ELDLLTDLVEGKGPLAQATMVQGVYLLKTPLPLPEYYIGLNVYFDKSGKLIYA  
 LDMSDTIGEGQKDAYGNPILNVDEDNEG YHTLAVATLADYEGLYIKDILNSSL  
 DKIKAIRQIPLAKYHRLGIFQAIRNAAAEADRLLPKTPKGYLNEVPNYRKKQM  
 EKNLKPVDYK TPIFNKALPNEKVDGDRAAKGHNINAETNNSVAVTPIRSEQQL  
 HKSQSDVNLPQTSSKNNFIYEILGYVSLCLLFLVTAGKKGKRARK\*

FIG. 1 CONTD

13 / 110

**Sequence description:**

- A] Length 3168 bp - 1056 aa (Partial sequence)
- B] Obvious signal peptide with Shine Dalgarno sequence upstream of the ATG start codon.

**ID-88**

**Clone RS-56**

GCAGGATACATCATGCACAAGCACGAGGCTATCGTGTCATGCTGGGGTCA  
ACCCAGGAAGACATGTCGGCACAAGCTGAAGATTTCTTTACAGTCTGTACA  
CAATAAAGAGACGGGTAAGAGCGCTTTTAATGACAAAGAACGACTAGCAA  
TT

AGYIMHKHEAIVSCWGQPRKTCRHLKISLQSVHNKETGKSAFNDKERLAI

**Sequence description:**

- A] Length: 153 bp - 51 aa (partial sequence)
- B] No signal peptide visible, insufficient sequence data to determine the presence of a Shine Dalgarno sequence.

**ID-89**

**Clone RS-58**

GTGTCATTTATGCAAAGAAAATCCTATTTAAAATCCATGAGTGTTCTTACT  
TTAACAGCTTGTCTTATATCAGGATATGTGGTTAAAGATATTGCTATGTTA  
CATGCAGTATCTGCCAGTGAGAAGAAAGCAAATAATGTCAGTCCGAGAGA  
AAATCTCTACAGGGCTGTCAATGATAATTGGCTAGCCAATACAAAACCTCA  
AACAAAGGGCAGACTAGTGTTAATAGTTTTTCAGAAATTGAGGATAAATTA  
AAGCAACTGTTAGTGTCTGATATGGCTAAAATGGCCTCAGGAAAGATTGA

**FIG. 1**<sub>CONT'D</sub>

14 / 110

AACAACCAATGATGAACAGAAAAAATGGTTGCATACTATAAACAAGGTA  
TGGACTTTTAAAACAAGAGATAAAAAATGGTCTCAAACCTCTAAAACCAGTT  
TTACAAAAACTTGAAGCAGTCTCTTCAATGAAAGACTTTCAAAGTTTGGCC  
CATGATTTTGTGATGAGTGGTTTTGTTTTACCATTTGGTTTGACTGTGGAAA  
CCAATGCTCGAGATAATAGCCAAAAGCAATTGGTGCTTCGTCAAGCACCC  
GCATTACTTGAATCACCTGACCAATATAAGAAGGGCAATAAAGAAGGTGA  
GGCTAAATTATCAGCTTACCGTACTTCAGCAATGGCTTTGCTTAAACAAGC  
TGGAAAAAGTAACATTGAAGATAGAAAAGTAAACAAGCTATAGCAT  
TTGATAGACTCTTATCAGAAAAAACGCAAGTTGATCAAAGTAAAATCACA  
GCTGAAAGTGAGACAGCTGCGGGGCGATATAACCCTGAAAGTATGGAAAC  
GGTTCACAATTACGCCAAGGAATTTGACTTTAAAGAATTGATTGAAAAACT  
AGTTGGGCCAACGAATAAGGCAGTCAATGTAGAAGATAAACTTATTTTA  
AACAGGTTAATGATGTTATAAATAGTAAACAATTAGCCAATATGAAAGCA  
TGGATGATGATTTCTATGCTAGTTGATCAATCAGATTTTCTAGGAGAACAA  
AATCGTCAAGCAGCGAGTGCTTTTAAAGAATGTTGCGTCTGGTTTGACTCAG  
ATTGAATCGAAAGAAAAAATGCTTACACCCAATTAG

MSFMQRKSYLKSM SVLTLTACLISGYVVKDIAM LHAV SASEKKANNVSPREN  
LYRAVNDNWLANTKLKQGQTSVNSFSEIEDKLKQLLVSDMAKMASGKIETTN  
DEQKKMVA YYKQGMDFKTRDKNGLKPLKPVLQKLEAVSSMKDFQSLAHDF  
VMSGFVLPFGLTVETNARDNSQKQLVLRQAPALLES PDQYKKGNKEGEAKLS  
AYRTSAMALLKQAGKSNIEDRKL VKQAI AFDRL LSEKTQVDQSKITAESETAA  
GRYNPESMETVHNYAKEFD FKELIEKL VGPTNKAVNVEDKTYFKQVNDVINS  
KQLANMKAWMMISMLVDQSDFLGEQNRQAASAFKNVASGLTQIESKEKMLT  
PN\*

Sequence description:

- A] Length: 1095 bp - 365 aa (full length gene)
- B] an GTG (possible ATG start codon located 7 bp further downstream) start codon with an obvious signal peptide. Shine Dalgarno sequence present upstream of the ORF.

ID-90

Clone RS-59

FIG. 1<sub>CONT'D</sub>

15 / 110

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAAAATGAG  
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC  
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC  
GTTATTTAAATATTTTTTATAG

MEMPKRNELLNKEIKMSIDKLRYPKEPESEHDKRPTFYLVVLILVTVAVILSLFK  
YFL\*

Sequence description:

- A] Length: 174 bp - 58 aa(full length gene)
- B] No obvious signal peptide, but Shine Dalgarno sequence is present upstream of ATG start codon.

ID-91

Clone RS-62 (partial sequence)

ATGCAGGTATTTTTAAATATTGTCAATAAATTCCTTTGATCCAGTTATTCATA  
TGGGTTTCGGGAGTTGTGATGCTAATTGTCATGACAGGTTTAGCCATGATAT  
TTGGAGTGAAGTTTTCTAAAGCACTTGAAGGTGGTAT

MQVFLNIVNKFFDPVIHMSGVVMLIVMTGLAMIFGVKFSKALEGG

Sequence description:

- A] Length:141 bp - 41 aa (partial sequence
- B] Shine Dalgarno sequence present upstream of ATG start codon with a possible signal peptide present

ID-92

FIG. 1<sub>CONT'D</sub>

16 / 110

Clone RS-69 (partial sequence)

ATGAAAAAGAAAACATTCAGTGCTTATAACTTTTTAACGGCTCTTATCCTT  
TGTCTTTTGACAGTGCTTTTATCTTCCATTTTATTGGATTATGACAGGAG  
CTTTTAA

MKKKTFSAYNFLTALILCLLTVLFIFPFYWIMTGAF

Sequence description:

- A) Length: 110 bp -36 aa (Partial sequence)
- B) Possible signal peptide with Shine Dalgarno sequence directly upstream of the ATG start codon.

ID-93

Clone RS-70

ATGACTGAGAACTGGTTACATACTAAAGATGGTTCAGATATTTATTATCGT  
GTCGTTGGTCAAGGTCAACCGATTGTTTTTTTACATGGCAATAGCTTAAGT  
AGTCGCTATTTTGATAAGCAAATAGCATATTTTCTAAGTATTACCAAGTT  
ATTGTTATGGATAGTAGAGGGCATGGCAAAAGTCATGCAAAGCTAAATAC  
CATTAGTTTCAGGCAAATAGCAGTTGACTTAAAGGATATCTTAGTTTCATTT  
AGAGATTGATAAAGTTATATTGGTAGGCCATAGCGATGGTGCTAATTTAGC  
TTTAGTTTTTCAAACGATGTTTCCAGATATGGTTAGAGGGCTTTTGCTTAAT  
TCAGGGAACCTGACTATTCATGGTCAGCGATGGTGGGATATTCTTTTAGTA  
AGGATTGCCTATAAATTCCTTCACTATTTAGGGAAACTCTTCCGTATATG  
AGGCAAAAAGCTCAAGTTATTTTCGCTTATGTTGGAGGATTTGAAGATTAGT  
CCAGCTGATTTACAGCATGTGTCAACTCCTGTAATGGTTTTGGTTGGAAAT  
AAGGACATAATTAAGTTAAATCATTCTAAGAAACTTGCTTCTTATTTTCCA  
AGGGGGGAGTTTTATTCTTTAGTTGGCTTTGGGCATCACATTATTAAGCAA  
GATTCCCATGTTTTTAATATTATTGCAAAAAGTTTATCAACGATACGTTG  
AAAGGAGAAATTGTTGAAAAAGCTAATTGA

MTENWLHTKDGSDIYYRVVGQGQPIVFLHGNSLSSRYFDKQIAYFSKYYQVIV  
MDSRGHGKSHAKLNTISFRQIAVDLKDILVHLEIDKVILVGHSDGANLALVFQ

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

18 / 110

Clone RS-73

TTGAGGGAACTTACTGGAAAATTTCAAGCGATTGCGATAAAATAAATCTT  
GCAGAGTTTTCTAGAGAAAGGAGGTCAGATTTATTGGAGTGGCAAGATCT  
AGCGCAGTTACCTGTATCTATTTTTAAAGACTATGTTACAGATGCTCAAGA  
CGCGGAAAAACCTTTTATATGGACAGAAGTATTTTAAAGGGAGATTAATCG  
CTCAAATCAAGAAATTATTTTGCATATTTGGCCGATGACTAAGACAGTCAT  
TCTGGGGATGTTAGATCGAGAATTACCACATTTAGAATTAGCTAAAAAAG  
AAATCATCAGTCGTGGTTATGAACCAGTTGTTCCGAATTTTGGAGGTCTCG  
CAGTTGTAGCTGATGAAGGAATTTTAAATTTTTCATTGGTTATTCCAGATGT  
TTTTGAGAGAAAATTGTCTATCTCAGATGGGTATCTTATAATGGTCGATTTT  
ATTAGAAGTATATTTTCGGATTTTATCAACCTATTGAGCACTTTGAAGTA  
GAGACCTCCTATTGTCCTGGTAAGTTTGATCTTAGTATAAATGGCAAAAAA  
TTTGCTGGCTTGGCTCAGCGCCGTATAAAGAATGGTATTGCGGTATCAATT  
TACCTTAGCGTTTGTGGCGATCAAAAAGGGCGGAGTCAAATGATTTTCAGAT  
TTTTATAAGATTGGTCTAGGTGATACGGGTAGTCCAATTGCTTATCCAAAT  
GTAGATCCTGAAATTATGGCTAATCTATCTGATCTATTAGATTGTCCTATG  
ACAGTAGAAGATGTTATTGATCGTATGTTGATTAGCCTTAAACAAGTAGGT  
TTTAATGATCGTTTACTGATGATTAGACCCGATTTAGTTGCAGAGTTTGAT  
AGATTTTCAGGCTAAGTCTATGGCTAATAAGGGGATGGTGAGCAGAGATGA  
ATAA

MRETYWKISSDCDKINLAEFRERRSDLLEWQDLAQLPVSIFKDYVTDAQDAE  
KPFIWTEVFLREINRSNQEIILHIWPMKTVILGMLDRELPHLELAKKEIISRGYE  
PVVRNFGGLAVVADEGILNFSLVIPDVFERKLSISDGYLIMVDFIRSIFSDFYQPI  
EHFEVETSYCPGKFDLSINGKKFAGLAQRRJKNIGIAVSIYLSVCGDQKGRSQMI  
SDFYKIGLGDGTGSPIA YPNVDPEIMANLSDLLDCPMTVEDVIDRMLISLKQVGF  
NDRLLMIRPDLVAEFDRFQAKSMANKGMVSRDE\*

Sequence description:

- A] Length: 921 bp -307 aa (Full-length gene sequence)
- B] No obvious Shine Dalgarno sequence upstream of the TTG start codon or signal peptide visible. Actual start point may be a further 85 bp downstream (TTG). This start point is preceded by a typical Shine-Dalgarno sequence.

FIG. 1<sub>CONT'D</sub>



20 / 110

## Clone RS-75

ATGACAACCTTACTACGAAGCTATAAACTGGAACGAAATTGAAGATGTTAT  
TGATAAATCAACTTGGGAAAACTAACCGAACAATTTTGGCTCGATACAC  
GTATCCCTTTATCAAATGACTTAGACGATTGGCGCAAACCTTCCGCTCAAG  
AAAAAGATCTTGTTGGCAAGGTTTTTGGAGGCTTAACCCTACTTGATACCA  
TGCAATCAGAACTGGTGTGGAAGCTATTCGTGCCGATGTTTCGCACGCCTC  
ACGAAGAAGCTGTCTTAAACAATATTCAATTCATGGAATCTGTTACGCTA  
AATCTTATTCTTCAATTTTCTCAACTTTAAATACTAAATCAGAAATTGAAG  
AAATTTTCGAGTGGACTAATAATAATGAGTTCCTTCAAGAAAAAGCACGT  
ATTATCAATGACATTTATGCTAATGGAAATGCCCTTCAAAAAAAGGTGGCT  
TCCACCTACCTCGAAACCTTCTTTTTTTATTCTGGCTTTTTTCACACCTCTTTA  
CTATTGGGAAATAATAAGTTAGCAAATGTTGCTGAAATCATTAAATTAAT  
TATTCGTGATGAATCTGTACATGGTACTTATATCGGTTACAAATTCAGCTT  
GGTTTTAACGAATTACCAGAAGATGAGCAAGAGAATTTTCGTGATTGGAT  
GTATGACCTCCTTTATCAGCTGTATGAAAACGAAGAAAAATACACCAAGA  
CACTTTATGATGGCGTAGGATGGACTGAAGAAGTTATGACCTTTTTACGCT  
ACAATGCTAATAAAGCTCTTATGAATTTAGGACAAGATCCTTTATCCAG  
ATACAGCAAATGATGTCAACCCAATTGTTATGAATGGTATTTCAACAGGAA  
CATCAAACCATGACTTCTTCTCTCAAGTAGGTAATGGTTACCTACTTGGTA  
GCGTTGAAGCTATGCATGATGACTATAACTATGGATTATAA

MTTYYEAINWNEIEDVIDKSTWEKLTEQFWLDTRIPLSNDLDDWRKLSAQEK  
DLVGKVFGLTLLDTMQSETGVEAIRADVTPHEEAVLNNIQFMESVHAKSY  
SSIFSTLNTKSEIEEIFEWTNNNEFLQEKARIINDIYANGNALQKKVASTYLETF  
LFYSGFFTPLYLGNKLANVAEIKLIIRDES VHGTYIGYKFQLGFNELPEDEQ  
ENFRDWMYDLLYQLYENEEKYTKTLYDGVGWTEEVMTFLRYNANKALMNL  
GQDPLFPDTANDVNPIVMNGISTGTSNHDFFSQVGNGYLLGSVEAMHDDYN  
YGL\*

## Sequence description:

- A) Length: 960 bp - 320 aa (full length gene)
- B) Shine Dalgarno sequence present upstream of  
ATG start codon, but no signal peptide  
present.

ID-98

FIG. 1 CONTD

21 / 110

Clone RS-77 (partial sequence)

ATGAATTGGTCACGTATCTGGGAACTCGTAAAAATTAATATCCTTTATTCA  
AACCTCAGACTCTATCGGCACTAAGAAAAAGCAAGAAAAGCATCCTAA  
AAAAGAATTTTCAGCTTATAAATCCATGTTTAGAAATCAGTTATTTAGAT  
TTTGCTCTTTTCAATAATTTATGTATTTCTCTTTGTATCACTTGATTTTAAAG  
AATATCCGGGCTATTTACGTTCTACATTGGTATCTTTACACTAGTATCCAT  
TATCTACTCTTTTATTGCGATGTACAGTGTTTTCTATGAGAGTGACGATGTT  
AA

MNWSRIWELVKINILYSNPQTLALRKKQEKHPKKEFSA YKSMFRNQLFQILL  
FSIIYVFLFVSLDFKEYPGYFTFYIGIFTLVSIIYSFIAMYSVFYESDDV

Sequence description:

- A) Length: 311 bp - 103 aa (Partial sequence)
- B) Shine Dalgarno sequence present upstream of  
ATG start codon, no obvious signal peptide at  
N-terminus.

ID-99

Clone RS-78 (partial sequence)

TAATCTTTTAGTCAACGGAGCAACAGGAAAATTGCAGGCTATGCGACAGA  
TATTCACCACATAATTTAGCAGAAGTCATTGATGCTGTCGTGTACATGAT  
TGATCACCCTAAAGCTAAATTAGATAAATTAATGGAATTTCTACCTGGTCC  
AGATTTTCCAACCTGGCGCTATCATTCAAGGAAAAGATGAAATTCGTAAGG  
CATATGAGACTGGTAAGGGGAGAGTAGCGGTTTCGCTCGCGAAGCTGCTATT  
GAAACCTTAAAAGGTGGTAAGAAACAAATTATTGTTACTGAAATTCCTTAT  
GAAGTTAAT

SFSQRSNRKLAGYATDIPPHNLAEVIDAVVYMIDHPKAKLDKLMEFLPGPDFPT  
GAIQKDEIRKAYETGKGRVAVRSRTAIETLKGKKQIIVTEIPYEVN

Sequence description:

- A) Length: 312 bp - 104 aa (Partial sequence)
- B) No obvious Shine Dalgarno sequence or a

FIG. 1 CONT'D

FIG. 1 CONT'D

FIG. 1 CONT'D

24 / 110

ID-103

Clone 2-11A

ATGGTATTTATGGCAAATAAGAAAAAAACAAAAGGAAAGAAAACCAGAA  
GACCTACTAAGGCAGAAATAGAGCGTCAAAGAGCTATTCAAAGGATGATT  
ACTGCTCTTGTTTTAACAATTATTCTCTTCTTTGGTATTATCAGATTAGGTA  
TTTTTGGTATTACAGTCTATAACGTCATCCGTTTTATGGTAGGTAGCTTGGC  
TACTTATTTATTGCGGCAACTTTAATCTACCTTTATTTCTTTAAATGGTTG  
CGAAAGAAAGATAGCTTAGTAGCAGGTTTTTTGATAGCTTCTTTAGGATTA  
TTGATTGAGTGGCATGCTTACCTTTTCTCAATGCCTATTTTGAAAGATAAA  
GAAATTTTGC GTTCAACTGCTCGATTAAATTGTGTCTGATTAAATGCAATTTA  
AAATCACTGTTTTTGCCGGTGGAGGTATGTTGGGTGCTTTGATTACAACC  
CAATTGCTTTTCTCTTTTCTAATATTGGTGCCTATATGATTGGTGTTCTCTTC  
ATCATTTTGGGTCTCTTTTAAATGAGTTCTCTGGAAGTTTATGACATCGTCG  
AATTTATTAGAGCTTTTAAAAATAAAGTGGCAGAGAAGCACGAGCAAAAT  
AAAAAGGAGCGTTTTGCTAAGCGAGAGATGAAAAAAGCAATCGCTGAACA  
AGAGCGCATAGAGCGTCAAAAAGCTGAAGAAGAAGCTTATTTAGCTTCGG  
TTAATGTAGACCCTGAAACGGGTGAGATTCTAGAGGATCAAGCTGAGGAC  
AATTTGGATGATGCGCTACCACCTGAGGTAAGTGAAACATCAACTCCGGT  
ATTTGAGCCAGAGATCCTTGCTTATGAGACATCGCCTCAAAATGATCCTTT  
ACCAGTAGAGCCGACAATTTATTTAGAAGACTATGATTGCGCCGATTCTTAA  
TATGAGAGAAAATGATGAGGAAATGTTTTATGATTTAGATGATGATGTAG  
ATGATAGTGATATAGAAAATGTGCACTTTACACCTAAAACGACACTGGTTT  
ATAAATTACCAACGATAGATTTATTTGCACCAGATAAGCCTAAAAATCAAT  
CCAAAGAAAAGGATTTAGTCCGAAAGAATATCAGAGTTTTAGAAGAAACA  
TTTAGAAGTTTTGGTATCGATGTAAAAGTAGAACGTGCTGAAATTTGACCA  
TCAGTTACTAAATATGAAATTAACACAGCAGTTGGAGTTGCTGTGAATCGT  
ATTTCAAATCTATCTGACGACCTAGCTCTTGCTCTTGACAGCAAAAGATGTG  
CGTATAGAAGCACCAATTCCTGGAAAATCATTAAATAGGTATTGAAGTTCCT  
AACTCAGAAAATTGCAACGGTTTTCTTTCCGCGAACTTTGGGAACAATCTGAT  
GCCAATCCTGAAAACCTTTTAGAAGTACCACTAGGAAAAGCTGTAAACGG  
CAATGCTCGCAGTTTTAACTTAGCTAGAATGCCGCATCTTTTGGTAGCTGG  
TTCAACTGGTTCAGGTAAATCTGTGGCAGTTAATGGAATTATTTCAAGTAT  
TTTGATGAAGGCACGTCCAGATCAAGTTAAGTTTATGATGATTGATCCCAA  
AATGGTTGAATTATCTGTTTATAATGATATTCCACATTTATTAATCCCTGTT  
GTAACCAATCCGCGTAAAGCAAGTAAGGCACTCCAAAAAGTTGTTGATGA  
AATGGAAAATCGATACGAGTTATTTAGCAAAATTTGGTGTGCGTAATATAG  
CAGGTTATAATACAAAGGTTGAAGAGTTTAAATGCTTCCTCTGAGCAAAAAC  
AAATGCCTTTGCCTTTAATCGTTGTCATTGTAGATGAATTGGCTGACTTGAT  
GATGGTTGCTAGTAAAGAAGTTGAAGATGCTATTATTCGTTTTGGGGCAAAA  
AGCACGTGCTGCAGGTATCCATATGATTCTTGCAACTCAACGTCATCCGT

FIG. 1 CONT'D

FIG. 1 CONTD

26 / 110

ATGTCACAAGAGCAAGGAAAAATTTATATTGTAGAAGATGATATGACGAT  
TGTGTCACCTTTTAAAAGATCATTTATCAGCTAGCTATCATGTCTCTAGTGC  
AGCAATTTTCGTGATGTGAAACAAGAAATTATCGCATTTCAACCCGATTTG  
ATACTAATGGATATTACGTTACCCTATTTTAATGGTTTTTACTGGACTGCAG  
AATTGCGTAAGTTTTTAACAATTCCTATTATTTTCATTTTCATCTAGTAATGA  
TGAAATGGATATGGTTATGGCATTAAATATGGGGGGTGATGACTTTATTTT  
AAAACCATTTCTCTAGCTGTATTAGATGCTAAGCTAACTGCTATTTTAAAG  
GAGAAGTCAACAATTTATCCAACAGGAATTAACCTTTTGGGGGATTTACGTT  
GACAAGAGAAGGGTTATTGTCTAGCCAAGATAAAGAGGTTATTTTATCGC  
CAACAGAAAATAAAATCCTATCTATCTTGCTCATGCATCCTAAACAAGTAG  
TCTCAAAAGAGTCTCTATTAGAGAACTTTGGGAAAATGATAGTTTTATTG  
ATCAAAATACACTTAATGTTAATATGACACGCTTACGTAAAAAAATTGTCC  
CAATAGGTTTTGATTACATTCATACAGTGAGAGGAGTTGGGTATTTACTAC  
AATGA

MSQEQGKIYIVEDDMTIVSLLKDHLSASYHVSSVSNFRDVKQEIIFQPDILM  
DITLPYFNGFYWTAELRKFLTPIIFISSNDEMMDVMALNMGGDDFISKPFSLA  
VLDAKLTAILRRSQFIQQLTFGGFTLTREGLLSSQDKEVILSPTENKILSILLM  
HPKQVVSKESSLLEKLWENDSFIDQNTLVNMTRLRKKIVPIGFDYIHTVRGVG  
YLLQ\*

Sequence description:

- A] Length: 669 bp - 223 aa (full-length gene sequence)
- B] Shine Dalgarno sequence present upstream of a GTG start codon. Was not identified directly by LEEP. This gene was found upstream of gene ID-10 described in WO 00/06736.

ID-105

Clone 2-20

ATGTATCAAACCTCAGACAAATAAGGAAAAATTTGTTTTATTTTTGAAATTA  
TTTATCCCAGTATTGATTTATCAATTTGCTAATTTTTCAGCTACTTTTATTGA  
TTCGGTTATGACTGGACAGTATAGTCAGCTACATTTGGCAGGTGTGTCAAC  
TGCTAGTAATTTATGGACTCCGTTTTTCGCTTTATTAGTAGGTATGATTTCA  
GCATTAGTACCAGTAGTTGGTCAACATTTGGGTAGAGGAAATAAAGAACA  
AATTCGCACAGAATTTTCATCAATTTCTATATTTAGGTTTGATACTGTCCTTA  
ATATTATTTTAAATCATGCAATTTATTGCTCAACCTGTCTTGGGGAGTTTGG

FIG. 1 CONT'D

FIG. 1 CONT'D

Clone 2-4A

TTGCTAGTTTCTTCTCTAGTTTCTTGTTTCATTTTTCTTGTCATTTTCGTCGTT  
 GTCTTCATCAACACGAAATAAGTCTATAAACTTATCAAATAATTTTCATAGA  
 CTTATTATATCAATTTTCAATAAAATGCTATAATAAAACCATGTCATTTTCA  
 TAAAAATTAGAAATCCATACGGTGAACATACCGTTAAAGAACTCCTTGA  
 AGATTATTTTTTGATTCCACGTAAGATTAGACATTTTTTGCGTGTTAAAAAA  
 CATGTACTTATAAACAATGAATTCATTAATTGGCAAACGTGTCGTCCAAGAA  
 AACGATACTATTACCTTAATCTTTGATGATGAGGATTACCCTACTAAAAAA  
 ATTCCTCTGGGCAGAGCAGAGCTTATTGATTGTCTTTATGAGGATGAACAT  
 CTTATTATCGTTAATAAACCTGAAGGTATGAAAACCTCACGGTAACCAACCA  
 AATGAAATAGCACTGTTAAATCATGTATCTGCCTATTCTGGACAAACATGC  
 TATGTTGTTTCATCGCCTAGATATGGAGACCAGTGGAGCTGTTTTATTGCT  
 AAAAATCCATTTATACTTCCCCTTATCAATCAACGCTTAGAACGAAAAGAA  
 ATTTGGCGTGAATATTGGGCTTTAGTTGAAGGAAAATTTTACCTAAGCAT  
 CAAGTTTTGAGAGACAAAATTGGACGGAACCGTCATGACAGACGTAAACG  
 AATCATTGATTCTAAAAACGGTCAACATGCTATGACAATCATTGACGTTTT  
 GAAGTATATCCAAAATAGTAGTCTCATAAAATGCCGACTGGAAACCGGAA  
 GAACCCATCAAATTCGCATTCACTTATCTCATCACGGACATCCTTTAATAG  
 GAGATCCCCTCTACAACCTTCTTCTAATAATGAAAGGTTAATGCTACACG  
 CTCACCGATTGACTCTATCCCATCCATTAACCTTGCAGAACTATTAGCGTAG  
 AGGCCCTTCATCTACTTTCGAGAAGGTTTTAAACAATTATAAAAAAGGAG  
 TTGGATAA

MLVSSLVSCSFFLVISLSSSTRNKSINLSNNFIDLLYQFSIKCYNKTMSFSLKIR  
 NPYGEHTVKELLEDYFLIPRKIRHFLRVKKHVLNNNEFINWQTVVQENDTITLIF  
 DDEDYPTKKIPLGRAELIDCLYEDEHLIIVNKPEGMKTHGNQPNELLNHVSA  
 YSGQTCYVVHRLDMETSGAVLFAKNPFIPLINQRLERKEIWREYWALVEGKF  
 SPKHQVLRDKIGRNRHDDRKRUIDSKNGQHAMTIIDVLKYIQNSSLIKCRLETG  
 RTHQIRIHLSHHGHPLIGDPLYNPSSNNERLMLHAHRLTLSHPLTCETISVEAPS  
 STFVKVLNNYKKGVG\*

Sequence description:

- A) Length: 1029 bp - 343 aa (Full length gene sequence)
- B) No obvious Shine-Dalgarno sequence upstream of the putative TTG start codon. Possesses a potential leader peptide sequence.

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

30 / 110

AAATAGTTTTACTTTAAATTACAAAGTTTTTAATTGGAGTTTTCTTAGTCAA  
AATACAGAGAAGCAAGGCACTTTATGGGAGAAAATGGCAGCAAATTGGCA  
TGTTTTGTTTAAATTTTATTTATGA

ELNATQPNRRTTYIIPESHISIAEQQRFLIESKGSSVALLNSDEFKRTAGEDRGF  
ERDKLRSLDIIPKGDLSNIGNTDIASQISLGFKKNAMQEHHLTKTFSQKDG  
KLSSVIEGMLAIGKEKVEKEIKYSGNLWQKLKAKAHCLVCCVDNLNFEDIKS  
YFQYYCHLNHQLKLPKGAILSaktevyrGGDFGRKNKDNVFGYRIPSLKTQ  
KGTLLAGADERIEQACDWGNIGMVIRSEDDGVTWGKRETIVNLRNPRVPL  
VTSGDYSGSPINMDMALVQDTSSKTKRIFSIYDMFPEGRGVISIANTPKEYTQI  
GGQSYLNLYNNGKKSUVFTIRDKGIVYNFKGKKTDYHVITETTKSDHSNLGDI  
YKKGQLLGNIFYTKHKTSPFRLAKSSYVWMSYSDDDGRTWSSPRDITASLRQ  
KGMKFLGIGPGKGIVLKWGPHAGRIIPAYSTNWKSHLRGSQSSRLIYSDDHG  
KTWHTGKAVNDNRILSNGEKIHSITMDNKKEQNTESVPVQLKNGDIKLFMRN  
LTGNLEVATSKDGETWQNHVKRYKEIHDAVYQLSAIRFEHDKKEYILLVNA  
NGPGKKCQDGYARLAQVNRNGSFKWLYHHIQDGSFA YNSVQQLNNDQFG  
VLYEHREKHQNSFTLNYKVFNWSFLSQNTEKQGTLEWKMAANWHVLFKFYL  
•

Sequence description:

- A) Length: 2052 bp - 684 aa (partial gene sequence)
- B) N-terminus has yet to be determined

ID-108

Clone 2-61

ATGCCTAAATTAATCGTATCTTTCCTCTGCATTTTATTATCCCTGACTTGTG  
TAAACTCTGTGCAAGCTGAAGAACATAAAGATATTATGCAAATTACCCGA  
GAAGCCGGATATGATGTTAAAGATATTAATAAACCTAAAGCGTCTATCGTT  
ATTGACAATAAAGGTCATATTTTGTGGGAAGATAACGCCGATTTAGAACGT  
GATCCCGCTAGCATGTCTAAAATGTTTACTTTATTTTACTATTTGAAGACT  
TAGCTAAAGGAAAAACAAACCTCAACACCACAGTGACTGCAACAGAAACA  
GACCAAGCCATAAGTAAGATTTATGAAATTAGTAATAACAATATTCATGCT  
GGGGTTGCTTATCCTATTCGTGAAGTACTATGACGGCTGTCCCGTCA  
TCTAATGTAGCAACTATTATGATTGCTAACCACTTATCACAAAACAATCCT  
GACGCCTTTATTAAACGAATCAATGAAACCGCCAAGAACTCGGTATGAC  
AAAAACTCACTTTTATAACCCCAAGTGGGGCGGTAGCGAGTGCTTTTAATGG  
ACTTTACTCCCAAAAGAATACGATAACAATGCTACTAACGTTACGACTGC

FIG. 1 CONT'D

31 / 110

ACGTGATCTATCAATTTTAACTATCATTTTCCTTAAAAAATACCCTGATATA  
CTGAACTATACAAAATATCCTGAAGTCAAGGCCATGGTCGGAACCTCTTAT  
GAAGAAACATTTACAACCTTATACTACTCTACCCCCGGCGCTAAATTTGGA  
TTAGAAGGAGTAGATGGCTTAAAAACTGGTTCTAGCCCTAGCGCTGCTTTT  
AATGCCTTAGTTACAGCTAAACGCCAGAATACTCGCTTGATAACTGTGGTT  
TTAGGAGTTGGCGATTGGTCAGACCAAGACGGAGAGTACTATCGTCATCC  
GTTTGTCAACGCTCTTGTAGAAAAAGGTTTTAAAGACGCTAAAAATATTTT  
TTCTAAAACTCCTGTATTAAAAGCCGTTAAACCTAAAAAAGAAGTTACTAA  
AACCAAACTAAATCTATTCAAGAACAGCCTCAAACAAAAGAACAGTGGT  
GGACAAAAACAGATCAATTTATCCAATCACATTTTGTATCTATTTTAATTG  
TTCTGGGCACCATCGCTAGCCTTTGTCTTTTAGCTGGGATAGTATTACTTAT  
AAAGCGCTCTAGATAA

MPKLIVSFLCILLSLTCVNSVQAEHDKDIMQITREAGYDVKDINKPKASIVIDN  
KGHILWEDNADLERDPASMSKMFTLYLLFEDLAKGKTNLNTTVTATETDQAI  
SKIYEISNNNIHAGVAYPIRELITMTAVPSSNVATIMIANHLSQNNPDAFIKRINE  
TAKKLGMTKTHFYNPSGAVASAFNGLYSPKEYDNNATNVTTARDLSILTYHF  
LKKYPDILNYTKYPEVKAMVGTPYEETFTTYNYSTPGAKFGLEGVDGLKTGS  
SPSAAFNALVTAKRQNTRLITVVLGVGDWSDQDGEYYRHPFVNALVEKGFK  
DAKNISSKTPVLKAVKPKKEVTKTKTKSIQEQPQTKEQWWTKTDQFIQSHFVS  
ILIVLGTIASLCLLAGIVLLIKRSR\*

Sequence description:

- A] Length: 1188 bp - 396 aa (full length gene)
- B] Shine Dalgarno sequence present upstream of  
ATG start codon, possesses a potential signal  
peptide

ID-109

Clone 45

ATGACTGAAAAATATTATAATTGGGCAACGCTTGGAACCGGCGTTATTGCC  
AACGAATTAGCCCAAGCACTGGAAGCACGTGGACAAAAATTATATTCTGT  
AGCTAATAGAACTTACGACAAAGGACTTGAATTTGCTAACAAATATGGTA  
TCCAAAAAGTTTATGATCACATAGATCAAGTATTTGAAGACCCTGAAGTGG  
ATATCATTTATATCTCTACTCCCCACAATACTCACATCTCATTTTTACGAAA

FIG. 1<sup>CONT'D</sup>

32 / 110

GGCTTTAGCAAATGGTAAGCACGTTCTTTGCGAAAAATCTATTACTTTAAA  
TAGTACTGAGCTTAAAGAAGCCATAGATTTAGCCGAAACTAACCATGTTGT  
CTTAGCTGAAGCCATGACTATTTTTCATATGCCAATTTACCGCCAATTTAAA  
AACATTAGTTGATAGTGGAATAATTAGGACCGTTAAAAATGATTCAAATGA  
ATTCGGAAGTTATAAAGAATATGATATGACTAACCGTTTTTTCAGTCGTG  
ACCTAGCAGGCGGTGCTTTGCTGGACATTGGTGTTTATGCACTTTCTTGAT  
TCGCTGTTTTATGTCAGAAGCACCTCACAAACATTACCTCTCAAGTTACATT  
TGCACCAACAGGGGTTGATGAACAAGTTGGTATCCTACTAACCAACCCAG  
CAAATGAGATGGCGACTGTCAGCCTTAGTTTACATGCAAAACAACCTAAA  
CGAGCAACTATCGCTTACGATAAAGGCTACATTGAACTTTTTGAATATCCG  
CGAGGACAAAAGGCAGTTATTACTTATACTGAGGATGGGCATCAAGATAT  
TATCGAAGCTGGCAAACTGAAAATGCTCTCCAATATGAGGTAGCTGATA  
TGGAAGAAGCCATTTAGGAAAACTAACACATGTACTTAACTATACC  
AAAGATGTTATGGATATCATGACACAGCTACGTCAAGAATGGGGATTTAC  
CTACCCAGAAGAAGAAAAATGA

MTEKYYNWATLGTGVIANELAQALEARGQKLYSVANRTYDKGLEFANKYGI  
QKVYDHIDQVFEDPEVDIIYISTPHNTHISFLRKALANGKHVLCESITLNSTEL  
KEAIDLAETNHVVLAEAMTIFHMPYRQLKTLVDSGKLGPLKMIQMNFGSYK  
EYDMTNRFFSRDLAGGALLDIGVYALSCIRWFMSEAPHNITSQVTFAPTGVDE  
QVGILLTNPANEMATVSLSLHAKQPKRATIAYDKGYIELFEYPRGQKAVITYT  
EDGHQDIIEAGKTENALQYEVADMEEAISGKTNHMYLNYTKDVMIMTQLR  
QEWGFTYPEEEK\*

Sequence description:

- A) Length: 984 bp - 328 aa (full length gene)
- B) Shine Dalgarno sequence present upstream of  
ATG start codon, possesses a potential signal  
peptide

ID-110

Clone 2-2

GTGTATTCTCCTGTAAATCTTCTAAAGGAAAAGTGATATTGTTAAAAAGT  
GATTTTCTAAAGAGCTTCATAGAAAGGAGAGGAAATATTTGTTTT

MYSPPVKSSKGKVILLKSDFLKSFIERRGNICF

FIG. 1 CONT'D

337 110

Sequence description:

- A] Length: 96 bp - 32 aa (partial sequence)
- B] GTG start codon - no obvious Shine-Dalgarno sequence
- Possesses a potential signal peptide

ID-111

Clone 2-3

AAATACTGTATCATTGCAACCTCAAATGCAGGTTTTGGAAACGAAGCATT  
ACAGGTGACAGCGATAAAGACTTGAAAATTATGGAACGAATTTCTCCATA  
TTCCGTCCAGAATTTCTAAATCGTTTCAATGGTGTTATTGAATTCTCTCAC  
CTAAGCAAAGATGACTTAAGCGAAATTGTAGATTTGATGCTTGATGAAGTT  
AACCAAACAATTGGCAAAAAAGGAATTGACCTTGTGGTAGATGAAAATGT  
TAAATCACACTTAATTGAACTGGGTTATGACGAAGCAATGGGAGTACGTC  
CATTGCGCCGTGTCATCGAGCAAGAAATTCGAGATCGCATCACAGACTACT  
ATCTCGATCATACAGACGTTAAACACCTAAAAGCTAATTTGCAAGATGGCC  
AAATCGTCATTTCTGAAAGATAA

KYCIIATSNAGFGNEAFTGDSKDLKIMERISPYFRPEFLNRFNGVIEFSLSKD  
DLSEIVDLMLDEVNQTIKKGIDLVVDENVKSHLIELGYDEAMGVRPLRRVIE  
QEIRDRTDYILDHTDVKHLKANLQDGQIVISER\*

Sequence description:

- A] Length: 429 bp - 143 aa (partial sequence)
- B] N-terminus yet to be elucidated. This gene was not in frame with nuc

ID-112

Clone 2-5

FIG. 1<sub>CONT'D</sub>

34 / 110

ATGTCAATGAATTTTTCATTTTACCACAATATTGGTCCTATTTTAATTATG  
 GTGTGATGGTAACCAATTATGATTTCAACATGTGTTGTTTTTTTGGAACTAT  
 TATAGGCGTGTTAATTGCTTTAGTAAAGCGTACTAATTTACATTTTCTCACA  
 ATATTAGCTAATTTCTATGTATGGGTATTTTCGTGGGACACCGATGGTAGTT  
 CAAATTATGATTGCTTTCGCATGGATGCATTTTAACAATTTACCAACAATT  
 AGCTTTGGTGTTTTAGATTTAGATTTTACACGACTTTTACCTGGTATCATT  
 TCATTTCTTAAATAGTGGTGCCTATATTTTCGGAAATTGTACGTGCAGGGA  
 TTGAGGCTGTACCATCTGGACAAATAGAAGCAGCTTACTCGTTGGGGATT  
 GACCTAAAAATACACTTCGCTATGTTATCTTACCCCAAGCTTTTAAAAATA  
 TTTTACCTGCTCTAGGGAATGAATTTATTACAATTATTAAAGATAGTGCTCT  
 CCTTCAAACATTGGTGTCATGGAATTATGGAACGGAGCACAATCAGTTGT  
 AACGGCTACTTACTCACCAGTTGCACCGTTATTATTTGCAGCATTTTACTAT  
 TTAATGTTGACAACGATTCTCTCAGCTTTGTTAAACAAATGGAGAAATAT  
 CTTGGGAAAGGGGTAAAAATAGATGGTTGA

MSMNFSFLPQYWSYFNYGVMVTIMISTCVVFFGTIIGVLIALVKRTNLHFLTIL  
 ANFYVWVFRGTPMVVQIMIAFAWMHFNNLPTISFGVLDLDFTRLLPGIIISLNS  
 GAYISEIVRAGIEAVPSGQIEAAYSLGIRPKNTLRYVILPQAFKNILPALGNEFITI  
 IKDSALLQTIGVMELWNGAQSVVTATYSPVAPLLFAAFYYLMLTTILSALLKQ  
 MEKYLKGKVKIDG\*

#### Sequence description:

- A) Length: 699 bp - 233 aa (full length gene)
- B) Shine-Dalgarno sequence preceded the 'ATG' start codon. Possesses a potential leader peptide sequence.

ID-113

Clone 2-7

ATGAAAGACCTATTACGAAATAGTCTAGAGCAAAGTGGAAATTTAAGTTT  
 TCAAGATATGATTTTACATATTCTTGTAGCAGCTTTATTGAGTGTAGTTATT  
 TATGTTTCCTATGCTTATACGCATAGTGGAACTGCCTATAGTAAAAAGTTT  
 AATGTTTCATTAATGACATTGACGGTCTTGACTGCAACAGTAATGACCGTT  
 ATTGGTAATAATGTAGCCTTGTCATTGGGTATGGTTCGGTGCCTTGTCAGTT  
 GTTCGTTTTAGGACAGCCATAAAAGATTCAAGAGATACAGTTTATATTTTT  
 TGGACCATAGTTGTTGGTATCTGTTGTGGTGTTCGGTGAAGTATGTGGTAGCT

FIG. 1 CONT'D

35 / 110

GCATTAGGAAGTAGCGTTATCTTTATCTTATTATGGGTTATGGGACGTGTT  
 AAAAACGAGAATCGTATGTTATTGATTGTGAAGTGCATAGAACACTAGA  
 AGTTGATTTAGAAGGAATTTTCTTCCAATATTTTGACGGAAAAGCTGTTCA  
 GCGTGTTAAAAATTCAACAATACTATTGAAATGATTTTCGAAATCTC  
 TAGAAAAGATTACGATAAGCAACTCCATGTAGATAATCAGTTAACTGAAA  
 AAGTGTACCAATTGGGAAATATTGATTATTTCAACATTGTTAGCCAAAGCG  
 ACGAAATCAATGGGTAG

MKDLLRNSLEQSGNLSFQDMILHILVAALLSVVIYVSYAYTHSGTAYSKKFNV  
 SLMTLTVLTATVMTVIGNNVALS LGMVGALSVVRFRTAIKDSRDTVYIFWTIV  
 VGICCGVGDYVVAALGSSVIFILLWVMGRVKNNRMLLIVKCDRTLEV DLEGI  
 FFQYFDGKAVQRVKNSTTNTIEMIFEISRKYDYDKQLHVDNQLTEKVYQLGNID  
 YFNIVSQSDEING\*

Sequence description:

- A) Length: 678 bp - 226 aa (full-length gene)
- B) ATG start codon is preceded by a Shine-Dalgarno sequence-Possesses a potential leader peptide sequence

ID-114

Clone 2-8

AAAAATTCATTTTAGATTTCATTTTACGACTATATACTCAGAAGTACCAAAC  
 CTAATCCAAGGTTTGAAAAAAGAAAGAAGGAAGTCAGTATGACAAACTAT  
 AAAACAACCTTTAAAGATGAGGCTATACGTGTTGAAGAGACAACAAAAGA  
 ATCATTTTACGATGTTGATATTGCCTTGTTTTAGCTGGTGGATCTATTTCA  
 GCAAAGTTCGCTCCTTATGCAGTAAAGTCTGGAGCAGTTGTAGTAGATAAC  
 ACGTCATATTTTCGTCAGAATCCTGATGTTCCACTAGTTGTTCCCTGAAGTAA  
 ATGCTCATGCCATGATTGGTCATAATGGTATCATAGCTTGTCCCAATTGTTT  
 TACTATTCAAATGATGATTGCTTTAGAGCCCATTCGTCAAAAATGGGGGAT  
 AGAGCGTGTTATAGTTTCCACCTATCAAGCTGTTTCGGGTTACAGGTGCACG  
 TGCTGTTGAAGAACTAAGGAACAGTTGAGACAAGTTTT

KFILD SFYDYILRSTKPNPRFEKRKKEVSMTNYKNNFKDEAIRVEETTKESFYD  
 VDIALFSAGGSISAKFAPYAVKSGAVVVDNTSYFRQNPDPVPLVVPEVNAHAMI  
 GHNGIACPN CSTIQMMIALEPIRQKWGIERVIVSTYQAVSGSGARAVEETKEQ  
 LRQV

FIG. 1<sub>CONT'D</sub>

36 / 110

Sequence description:

- A] Length: 499 bp - 165 aa (partial sequence)
- B] N-terminus has yet to be determined

ID-115

Clone 2-9

ATGACAAATGAATTGATAATGCAAGCTTTTGAGTGGTATTTACCTAGTGAT  
GGGAATCACTGGAAGAAATTAGAGGAGTCTATATCAGACCTTAAAAAACT  
TGGAATTAGTAAAATCTGGTTACCACCAGCATTTAAGGGAAGTAGCAGTG  
ATGATGTAGGATATGGTGTTTATGATCTCTTTGATTTAGGAGAATTTGACC  
AGAATGGAACAATTAGAACAAAATATGGTAGGAAAGAAGAGTATCTAAA  
GCTTATTAAGTCGTTAAAGGCAAATGGCATTAAACCGTTTGCAGATATCGT  
TCTTAACCATAAAGCCAATGGTGATCATAAAGAAAAATTTCAAGTCATCA  
AAGTCAATCCTGAAAATCGTCAAGAAGCATTAAAGTGAACCCTATGAGATT  
GAAGGATGGACGGGATTTGATTTCCCAGGTAGACAGGGTGAGTACAATGA  
TTTT

MTNELIMQAFEWYLPDGNHWKKLEESISDLKKLGISKIWLPPAFKGTSSDDV  
GYGVYDLFDLGEFDQNGTIRTKYGRKEEYLLIKSLKANGIKPFADIVLNHKA  
NGDHKEKFQVIKVNPNRQEALSEPYEIEGWTGDFPGRQGEYNDF

Sequence description:

- A] Length: 456 bp - 152 aa (partial sequence)
- B] ATG start codon is preceded by a Shine-Dalgarno sequence, no leader peptide sequence.

ID-116

Clone 2-10

FIG. 1<sub>CONT'D</sub>

57 / 110

ATGGAGGTTCTTATGAAGAAAAGTGTTAGTAAGTAGTCTTTTGGTTTTAGGG  
ATTACGATAACGTTACAACCAGTAGTTGAGGCTAAGGGGCCAAAAGTAGC  
TTATACACAAGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAG  
TCACTACTATTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGC  
CGATTACTGTTAGTTTTGATATTGATGATACTGCTTTTCAGTAGTCAATA  
TTTTCAATATGGTAAAGAATATGTAACCTCGGATCGTTTGATTTTCTTCAT  
AAACAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCAT  
TCCCAAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAG  
ATAAAATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAG  
GGCGAGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTTGTA  
CCATCTGAT

MEVLMKKVLVSSLLVLGITITLQPVVEAKGPKVA YTQEGMTALSDTNKDKVT  
TISIDEIQKSLEGKKPITVSFDIDTLLFSSQYFQYGYKEYVTPGSFDFLHKQKFW  
DLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGSMYKEGEVDKTA  
KALAKDFKFVPSD

Sequence description:

- A) Length: 516 bp - 172 aa (partial sequence)
- B) ATG start codon is preceded by a Shine-Dalgarno sequence, Possesses a leader peptide sequence.

ID-117

Clone 2-17

ATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGACAAAGATTAGTCGT  
CGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTATTATTGTCAGGATGT  
GTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAACTCCGAATGTTTCAGC  
GCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTAAATTCTTTTCGTAGTTT  
AGATCAGCTAACTAGCTTTAAAGAGATTCTTTGGGTTATTGGTCAAAATGT  
AGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGTTACTATCCCTAAAG  
CCAAGTTTACGGAAATATAAAAGCGTTATATTACTTGCTTTCTTGATGTCTC  
TTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTTAATAGATGCTAATCG  
GGTTTTTGAAATCGACGATCTATGGACAAATACCTTAGGCGGTCTTTTCGC  
CCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCTAACTATTAGAAAATG  
A

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

39 / 110

B) ATG start codon is preceded by an possible Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-119

Clone 3-7

ATGGAAAAAGAAAAAAATTAGGTCTTTTACCACTAACAATGCTTGTCATT  
GGCTCTCTTATCGGTGGCGGAATCTTTGATTAAATGCAAAATATGAGTTCC  
AGAGCCGGTTTGGTACCAATGCTTATTGCTTGGGTAACTACTGCTATCGGG  
ATGGGAACTTTCGTTTTAAGTTTTCAAAATTTATCTGAAAAAGGCCGGAC  
CTAACAGCTGGAATCTTTAGTTACGCTAAAGAGGGGTTTGGAACTTTATG  
GGATTTAACTCTGCATGGGGTTATTGGTTATCAGCTTGGCTTGGAAATGTT  
GCCTACGCTGCACTCTTATTCAGTTCACCTCGGTTATTTCTTTAAATTCTTTG  
GTAATGGAAATAATATCATCTCAATTATTGGAGCAAGTATAGTTATTTGGG  
TTGTCCATTTCTTAATTTTAAGAGGTGTTAATACAGCTGCATTTATTAATAC  
CGTAGTTACCTTTGCAAAATTAGTACCTGTTATTATTTTCTTAATTTACGCG  
TTATTAGCTTTCAAATTTAACATTTTATGCTTTGATATCTGGGGAAATGGAT  
TACATCAATCAATTTTCAACCAAGTCAATTCAACTATGAAAACCGCTGTTT  
GGGTATTTATTGGTATTGAGGGCGCCGTTGTCTTCTCAGGTCGTGCTAAAA  
AACACTCTGATATTGGTAAAGCAAGTATCCTAGCATTATTCACATATGATTT  
CACTTTATGTATTGATTTCTGTTTTATCACTTGGTATCATGTCACGTCCAGA  
ACTTGCAAACTTAAAAACACCAAGCTATGGCTTACGTTCTAGAAAAAGCTGT  
TGGTCACTGGGGTGCTATCTTAGTTAACCTTGGTGTTATCATTTCAGTATTT  
GGCGCTATTCTTGCTTGGACTTTATTTGCAGCAGAATTACCATATCAAGCT  
GCTAAAGAAGGTGCTTTTCTAAATTTTTTGCAAAAGAAAAATAAAAACAA  
AGCTCCAATCAACTCACTCTTAGTCACTAATCTTTGTGTACAAGCATTCTTA  
ATCACGTTCTTATTCACACAAAGTGCTTATCGTTTTGGTTTCGCATTAGCAT  
CATCTGCTATCTTAATTCCTTATGCTTTTACAGCACTATATCAATTACAATT  
CACACTCCGTGAGGATAAGTCAACTCCAGGACATCAAAGAATTTAATTA  
TCGGTATCCTCGCTACAATCTATGCTGTTTACCTTATCTACGCTGGTGGTTT  
TGATTACTTACTTTTGACAATGATTGCTTATACTCTAGGTATGATTCTCTAT  
ATTAAAATGAGAAAAGATGACAAGCTTGGCGTAATCATGGTCATAGCTGT  
TTCCAGTGTGAAATTGTTATCC

MEKEKKLGLLPLTMLVIGSLIGGGIFDLMQNMSSRAGLVPMLIAWVITAIGMG  
TFVLSFQNLSEKRPDLTAGIFS YAKEGFGNFMGFNSAWGYWLSAWLGNVAY  
AALLFSSLGYFFKFFGNGNNIISIIGASIVIWVVFHLILRGVNNTAAFINTVVF  
LVPVHIFLISALLAFKFNIFSLDIWGNGLHQSFNQVNSTMKTAVWVFIGIEGAV

FIG. 1 CONT'D

40 / 110

VFSGRAKKHSDIGKASILALFTMISLYVLISVLSLGIMSRPELANLKTPAMAYV  
LEKAVGHWGAILVNLGVIIISVFGAILAWTLFAAELPYQAAKEGAFPKFFAKEN  
KNKAPINSLLVTNLCVQAFLITFLTQSA YRFGFALASSAILIPYAFTALYQLQF  
TLREDKSTPGHQKNLIIGILATIYAVYLIYAGGFDYLLLTMIA YTLGMILYIKMR  
KDDKLGVIMVIAVSSVKLLS

Sequence description:

- A] Length: 1356 bp - 452 aa (partial sequence)
- B] ATG start codon is preceded by an possible Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.

ID-120

Clone 3-8

ATGAAATTTGAAAAACGGCAGGTCTATTATGTTGTCATAACATTTGCTATT  
TGCTATGCTATACAGGCTTATTGGGGAGCTGTTTCTAATATTTAACTACGC  
TTCATAAGGCAATATTTCTTTTTTGGATGGGAGCTGGAATTGCCTATATTAT  
TAATATTGTAATGTCAGTCTATGAGCGATTATATATAAAGCTTTTTAAAGG  
ATCTAGACTATTAATGGCAATCAAGCGTAGTGTTTCTATGATTTTATCCTAT  
GCAACTTTTATTGGTTTAATTGTCTGGCTATTTTCAATTGTCATTCCAGATT  
TGATTTCTAGTTTGAGTTCTTTATTGGTTATTGATACCGGAGCACTTGCTAA  
ATTGGTTAATAATCTCAATGAAAATAAACAATTTCTGAGGCTTTAAATTA  
TATGGGAACAGATAAAGACTTAGTTTCTACTTTAAGTGGTTATAGCCAGCA  
GATTTTGAAGCAAGTTTTATCTGTTTTAACAATTTACTAACCTCAGTTTCC  
TCTATTGCGGCAACACTTCTGAATGTTTTGTTAGTTTTATTTTTTCAATTA  
CGTTTTGGCAAACAAGGAGCAGTTGGGACGTCAATTTAATTGTTAATTGA  
TACCTATTTAGGTTCAACAGGCAAAACATTCCATTACGTTTCGTCATATCCTT  
CATCAACGTTTCCATGGTTTTTTTTGTAAGCCAAACTTTAGAAGCTATGATT  
TAGGAAGTTTGACGGTTATTGGTATGTTGATCTTCCAATTTCTTTATGCTTT  
AACAGTTGGGGTTTTAGTTGCTTTTACAGCTCTAATACCGGTTGTGGGAGC  
CTACATTGGTGTTACAATCGGTTTCATCTTAATTGCTACTGAATCGCTTACT  
GAAGCATTCTTGTTGTTCTTTCTTGATCCTTTTACAACAATTTGAGGGAA  
ATGTCATTTATCCGAAAGTTGTCGGTGGATCGATTGGACTGCCTTCTATGT  
GGGTTTTAATGGCTATTACTATCGGAGGTGCTTTATGGGGGATCTTAGGCA

FIG. 1<sub>CONTD</sub>

41 / 110

TGTTACTTGCTGTTCTGTTGCAGCTACTATCTATCAGATTGTAAAAGATCA  
TATTATCAAGCGACAAACGCTTAGAAATCGTGCACGAACCTATCGTTAA

MKFEKRQVYYVVITFAICYAIQAYWGAVSNILTTLHKAIFPFLMGAGIAYIINI  
VMSVYERLYIKLFGKSRLMAIKRSVSMILSYATFIGLIVWLFSIVIPDLISLSS  
LLVIDTGALAKLVNNLNENKQISEALNYMGTDKDLVSTLSGYSQQILKQVLSV  
LTNLLTSVSSIAATLLNVFVSFIFSIIYVLANKQLGRQFNLLIDTYLGSTGKTFH  
YVRHILHQRFHGFFVSQTLEAMILGSLTVIGMLIFQFPYALTVGVLVAFTALIP  
VVGAYIGVTIGFILATESLTEAFLFVLILLQQFEGNVIYPKVVGGSIGLPSM  
WVLMAITIGGALWGILGMLLAVPVAATIIYQIVKDHIKRQTLRNRARTYR\*

Sequence description:

- A) Length: 1134 bp - 378 aa (full-length gene)
- B) ATG start codon is preceded by an typical Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.

ID-121

Identical to ID-68, as described in WO 00/06736

ID-122

Clone 3-16

GTGATTACAATTAAAAAGGAATCTGTTATCAAACCTATTGAAGTATGCTTTT  
GGCATTATAATGGGATTTATTATCTTAGCTATTGTAATAGGTGGGCTCCTA  
TTTGCATACTACGTTAGTCGTTCTCCGAAATTAACCGATCAAGCTTTAAAA  
TCCGTTAACTCTAGTTTGGTTTATGATGGTAATAATAAACTTATTGCCGATT  
TAGGCTCAGAAAAGCGTGAAAGTGTTAGTGCGGATAGCATTCCACTAAAT  
TTGGTTAACGCTATCACTTCTATAGAAGATAAACGTTTCTTTAAACATAGA  
GGTGTCGATATTTATCGTATTTTAGGTGCAGCTTGGCATAACCTTGTTAGTA  
GTAATACGCAAGGTGGTTCAACCCTTGATCAACAGTTGATTAACTGGCTT  
ACTTTTCTACCAATAAATCTGACCAAACGTTAAACGTAAATCACAGGAA  
GTTTGGCTTGCGCTTCAAATGGAGCGTAAATACACCAAAGAAGAAATTCTT  
ACTTTCTATATTAATAAAAGTTTATATGGGAAATGGGAATTATGGTATGAGA

FIG. 1 CONTD

ACAAACAGCTAAATCATACTTTGGTAAAGACCTAAAGGAATTATCTATTGCA  
CAACTTGCTTTGCTCGCTGGTATTCCTCAAGCACCTACACAATATGACCCTT  
ATAAAAAACCCAGAATCTGCTCAAACAAGACGTAATACCGTTCTTCAGCAG  
ATGTATCAAGATAAAAAACATTTCTAAAAAGGAATACGACCAAGCTGTTGC  
AACTCCAGTAACTGATGGCTTAAAGAATTAAAGCAAAAATCTACTTATCC  
AAAATATATGGATAACTACTTAAAAACAAGTTATTAGTGAAGTTAAACAAA  
AAACTGGTAAAGATATCTTTACTGCTGGGCTAAAAGTGTATACTAATATCA  
ACACTGATGCACAAAAACAATATATGACATCTACAACAGTGATACTTAC  
ATCGCTTATCCAAACAATGAATTACAAATAGCATCTACCATCATGGATGCG  
ACTAATGGTAAAGTCATTGCACAATTAGGCGGGCGTCATCAGAAATGAAAA  
TATTTTCAATTTGGGACAAATCAATCTGTCTTAAACAGACCGCGATTGGGGTTC  
TACAATGAAACCTATCTCAGCTTATGCACCTGCTATTGATAGTGGTGTCTA  
TAATTTCAACAGGTCAATCATTAAACGACTCAGTTTACTACTGGCCTGGTAC  
TTCTACTCAACTATATGACTGGGATCGTCAATATATGGGTTGGATGAGTAT  
GCAGACCGCTATTCAACAATCAGCTAACGTCCCTGCTGTGAGAGCACTTGA  
AGCCGCTGGATTAGACGAAGCAAAATCTTTCCTTGAAAAATTAGGCATAT  
ACTATCCAGAAATG

MITIKKESVIKLLKYAFGIIMGFIILAIVIGGLLFAYYVSRSPKLTQALKSVNSS  
LVYDGNKLIADLGSEKRESVSADSIPLNLVNAITSIEDKRFFKHGVDIYRILG  
AAWHNLVSSNTQGGSTLDQQLIKLAYFSTNKSQDTLKRKSQEVWLALQMER  
KYTKEEILTFYINKVYMGNGNYGMRTTAKSYFGKDLKELSIAQLALLAGIPQA  
PTQYDPYKNPESAQTRRNTVLQQMYQDKNISKKEYDQAVATPVTDLKELK  
QKSTYPKYMDNYLKQVISEVKQKTGKDIFTAGLKVYTNINTDAQQLYDIYN  
SDTYIAYPNNELQIASTIMDATNGKVIAQLGGRHQENISFGTNQSVLTD RDW  
GSTMKPISAYAPAI DSGVYNSTGQSLNDSVYYWPGTSTQLYDWD RQYMGWM  
SMQTAIQSRNVPAVRALEAAGLDEAKSFLEKLGIIYPPEM

**Sequence description:**

- A) Length: 1386 bp - 462 aa (partial sequence)  
B) GTG start codon is preceded by an  
typical Shine-Dalgarno sequence. Possesses a  
potential leader peptide sequence.

ID-123

### Clone 3-17

FIG. 1 CONT'D

43 / 110

ATGGCTAATGTATATGATTTAGCAAATGAATTAGAACGTGCTGTTCTGTGCT  
TTACCAGAATACCAAGCAGTTTTAACTGCAAAAGCAGCTATTGAAAATGA  
TGCGGATGCACAAGTGCTTTGGCAAGACTTTTTGGCTACCCAATCAAAAGT  
TCAAGAAATGATGCAATCTGGCCAAATGCCAAGTCAAGAAGAACAAGATG  
AAATGTCTAAACTTGGGGAAAAAATTGAATCCAATGACCTTTTAAAAGTTT  
ATTTTGACCAACAACAACGGTTGTCTGTCTATATGTCTGATATCGAAAAAA  
TTGTCTTTGCACCCATGCAGGACTTGATGTAA

MANVYDLANELERAVRALPEYQAVLTAKAAIENDADAQVLWQDFLATQSK  
VQEMMQSGQMPSQEEQDEMSKLGEKIESNDLLKVYFDQQRLSVYMSDIEKI  
VFAPMQDLM\*

Sequence description:

- A) Length: 336 bp - 112 aa (full length sequence)
- B) ATG start codon is preceded by an  
typical Shine-Dalgarno sequence. No obvious  
potential leader peptide sequence.

ID-124

Clone 3-26

ATGGCAGAAATCACAGCTAAACTTGTAAAAGAATTGCGTGAAAAATCAGG  
TGCAGGCGTTATGGACGCTAAAAAAGCATTAGTAGAAACTGATGGTGACC  
TTGATAAAGCGATTGAATTACTTCGCGAAAAAGGTATGGCTAAAGCAGCT  
AAAAAAGCAGACCGTGTTGCTGCTGAAGGTTTAACAGGTGTTTATGTTGAT  
GGTAACGTTGCAGCAGTTATTGAAGTTAA

MAEITAKLVKELREKSGAGVMDAKKALVETDGDLDKAIELLREKGMAKAAK  
KADRVAEGLTGVYVDGNVAIV

Sequence description:

- A) Length: 230 bp - 76 aa (partial sequence)
- B) ATG start codon is preceded by an  
typical Shine-Dalgarno sequence. No obvious  
potential leader peptide sequence.

FIG. 1<sub>CONTD</sub>

44 / 110

ID-125

Clone 3-33

ATGATAAAAAACCTGTTATTAACAGGTTTTTTATCATTTAATGACGGAAAA  
CTGGACACAAATTATTTTCTTGTATAATTAAATATATTATTTCTTATCAGG  
AGGTTATGATGACATTAGAGAAACGATTAA

MIKNLLLTGFLSFNDGKLDTNFYFSCIIFYIISYQEVMMTLEKRF

Sequence description:

- A] Length: 134 bp - 44 aa (partial sequence)
- B] ATG start codon is preceded by an  
typical Shine-Dalgarno sequence. Possible  
potential leader peptide sequence.

ID-126

Clone 3-41

ATGAAAAATAATAAAAAATAATGGTTTTCTGAAAAATTCCTTTATTTACATA  
TTATTGATTATTGCGGTTATTACAACCTTTCAATACTATTAA

MKNNKNNGFLKNSFIYILLIIVITTFQYYL

Sequence description:

- A] Length: 94 bp - 31 aa (partial sequence)
- B] ATG start codon is preceded by a  
possible Shine-Dalgarno sequence. Potential  
leader peptide sequence.

FIG. 1 CONT'D

45 / 110

ID-127

Clone 3-42

ATGTTAGATATTATCTTATCCGGAATTTTCGCAAGGATTACTTTGGTCAATTA  
TGGCAATTGGCGTGTTTATCACTTTTCGTATCTTAGACATAGCCGATCTCTC  
TGCAGAAGGGGCTTCCCTATGGGGGCTGCAGTTTGCGCCTTATGTATCGT  
TAA

MLDIILSGISQGLLWSIMAIGVFITFRILDIADLSAEGAFPMGAAVCALCIV

Sequence description:

- A] Length: 158 bp - 52 aa (partial sequence)
- B] ATG start codon is preceded by a possible Shine-Dalgarno sequence. Potential leader peptide sequence.

ID-128

Clone 3-43

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAATGAG  
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC  
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC  
GTTATTAA

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLF

Sequence description:

- A] Length: 161 bp - 53 aa (full-length gene)
- B] ATG start codon is preceded by a possible Shine-Dalgarno sequence. Potential

FIG. 1<sub>CONT'D</sub>

46 / 110

leader peptide sequence.

ID-129

Clone 3-44

GTGGTAAGTAAATTGAGTTTAACAACGATTTTTGCATTGCTATTTTCATCA  
ATGCTAATTTACGCAACACCTCTTATCTTTACAAGTATTGGGGGAACCTTC  
TCTGAACGTGGTGGTATCGTCAACGTTGGTTTAGAAGGAATTATGGTAATT  
GGAGCTTTCTCAGGCGTTGTATTTAA

MVSKLSLTTIFALLFSSMLIYATPLIFTSIGGTFSERGGIVNVGLEGIMVIGAFSG  
VVF

Sequence description:

- A) Length: 179 bp - 59 aa (partial sequence)
- B) GTG start codon is preceded by a possible Shine-Dalgarno sequence. Potential leader peptide sequence.

ID-130

Clone 3-46/47

ATGAGAATTATTGCAATAACTGAAAAGGTTATAAAAAGAACTGTTTCGTGAT  
AAAAGAACACTTGCTATGATGTTTTTAGCACCTATTTAATTATGTTTTGA  
TGAATGTTATGTTTTCTGCGAATAGTAATACAAAAGTTAAGATTGGAACCTA  
TTAACGTTAACACGAAGGTCGTTTCAAATTTAGATAATATTAAGCATATTC  
AAGTGAGATCATTAAATTTAACTCATCTGCTAAAAAAGCACTCAAATCAA  
ATAAAATTGATGCTCTTATTTTCGGAGGACAATAAATCTTATACTGTCTTCT  
ATGCGAATACAGATTCTTCAAAGACGACTTTAACAAGACAAGCTTTTAAA  
ACCGCTGTTAATACAATGAACAGTAAGGAACTGATTTTCGCAAGTTAAATTT  
TTAGCTAATAAGAATCCGAACTAGCACAATCCTTACAACTCGCTCCAAA  
TATATCAAAGAAAAATATAATTACGGAAATAAAAATACAGGCTTTTTTGC  
AAAAATGATACCAATACTAATGGGATTTATGGTCTTCTTCTTGGTTTTT

FIG. 1 CONT'D

47 / 110

MRIIAITEKVIKELFRDKRTLAMMFLAPILIMFLMNVMFSA NSNTKVKIGTINV  
NTKVVS NLDNIKHQVRSFKFNSSAKKALKSNKIDALISEDNKSYTVFYANTDS  
SKTTLTRQAFKTA VNTMNSKELISQVKILANKNP KLAQSLQTRSKYIKEKYNY  
GNKNTGFFAKMIPILMGFMVFFLVF

Sequence description:

- A) Length: 558 bp - 186 aa (partial sequence)
- B) ATG start codon is preceded by a possible Shine-Dalgarno sequence. Potential leader peptide sequence. C-terminus has yet to be determined.

ID-131

Clone 3-48

GTGATTATCGTTATGAGTAAACATCAAGAAATTTTGGAGTACCTAGAAAAT  
TTAGCTGTTGGTAAGAGGGTTAGTGTACGCAGTATTTCAAATCATTAA

MIIVMSKHQEILEYLENLAVGKRVS VRSISNHL

Sequence description:

- A) Length: 100 bp - 33 aa (partial sequence)
- B) GTG start codon is not preceded by a obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-132

Clone 2-c53

FIG. 1 CONT'D

48 / 110

ATGTATAGAGAAATTACCGCTGTCGAACACGATCGCTTTGTGAGCGAATCC  
AACCAAACAAACCTACTTCAATCTCTTAATTGGCCCAAAGTAAAAGACAA  
CTGGGGTAGTCAATTACTTGGCTTTTTTGGCGGTGAAACCCAAATTGCCAG  
CGCTAGTATTCTCATCAAATCACTTCCTCTTGGCTTCTCCATGCTGTATATT  
CCGCGTGGACCAATCATGGATTACTCCAATCTAGATATTGTAACCTAAGGTC  
CTTAAGGACCTTAAAGCTTTTGGCAAAAAACAAAGAGCTCTCTTTATCAAG  
TGTGATCCTCTCATCTATTT

MYREITAVEHDFVSESNTNLLQSLNWPVKVDNWGSQLLGFFDGETQIASA  
SILIKSLPLGFSMLYIPRGPIMDYSLNDIVTKVLKDLKAFGKKQRALFIKCDPLI  
Y

Sequence description:

- A] Length: 326 bp - 108 aa (partial sequence)
- B] ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-133

Clone 2-c59

ATGGACAAGAAAAAATCTTAGTAACGGGTATTGTGCCTAAAGAAGGTCT  
AAGAAAGCTTATGGACCGATTGTGTTACTTATTCAGAAGATCGCCCAT  
TTCACGTGACTATGTGTTAGAGCATTTATCTGAATATGACGGATGGTTACT  
CATGGGACAAAAAGGTGATAAAGAGATGATTGATGCAGGTGAAAACCTAC  
AAATTATTTCTTT

MDKKKILVTGIVPKEGLRKLMDRFDVTYSEDRPFSRDYVLEHLSEYDGWLLM  
GQKGDKEMIDAGENLQIIS

Sequence description:

- A] Length: 215 bp - 71 aa (partial sequence)

FIG. 1 CONT'D

49 / 110

B) ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-134

Clone 2-c62

ATTTCGAAAGATGACTACCAAAATATTAGTTTTGGACAGGATCCAGAAGTT  
GTTGATTATGCTGGTCTGTTTGAAAAACGCCGTCCAGTTTTAGAAAAAGCA  
GTTAAAAATTTCTTGCAAGAAGAGAGCTACGAGAATGCTATCTGATTTT  
TTGCAAGAAGAAAAATGGGTAACCTGATTTTGCTGAATTTATGGCGATCAA  
AGAACATTTTGGTAATAAGGCGCTTCAAGAATGGGATGACAAGGCTATTA  
TACGCCGCGAAGAAGAAGCCTTAGCAGGATATCGTCAAAAGCTTAGTGAA  
GTGATAAAATATCATGAAGTAACGCAATATTTCTTTTACAAACAATGGTTT  
GAGTTAAAAGAATATGCTAATGATAAAGGGATTCAAATTATCGGTGATAT  
GCCAATCTACGTTTCTGCCGATAGTGTAGAAGTTTGGACAATGCCTGAACT  
GTTT

ISKDDYQNISFGQDPEVVDYAGLFEKRRPVLEKAVKNFLQEERATRMLSDFLQ  
EEKWVTDFAEFMAIKEHFGNKALQEWDDKAIIRREEEALAGYRQKLSEVIKY  
HEVTQYFFYKQWFELKEYANDKGIQIIGDMPYVSADSVEVWTMPFLF

A) Length: 459 bp - 153 aa (partial sequence)  
B) More sequencing is required to determine the  
N- and C-termini  
enzyme). - *Streptococcus pneumoniae* (63%)

ID-135

Identical to ID-108 described in WO 00/06736

Clone 2-c63

ID-136

Clone 2-c66

FIG. 1<sub>CONT'D</sub>

50 / 110

ATGGCAAAACAGAAAAATAACTGGCGCCGTGTTGGAGTTGGTGTCTTAC  
 ACTTGCTTCAGTTGCGACTCTTGCTGCATGTGGAAGTAAATCAGCTTCCCA  
 GGATTCTAATGGAGCGATTAATTGGGCTATTCCAACAGAAATCAATACACT  
 AGATTTATCTAAAGTTACAGACACTTACTCAAATCTAGCTATTGGTAACTC  
 TAGTAGTAATTTCTTCGCTTAGATAAAGATGGAAAGACAAGACCAGACTT  
 GGCTACTAAAGTTGATGTTTCAAAAGATGGCTTAACTTATACAGCTACATT  
 ACGTAAAGGCTTGAAGTGGTCAGATGGCAGTAACTTACTGCAAAGGATT  
 TTGTTTATTTCATGGCAACGTTTAGTTGATCCTAAAACAGCTTCACAATATG  
 CTTACCTTGCTGTTGAAGGGCATGTGCTTAATGCCGATAAAATCAACGAAG  
 GACAAGAGAAAGACTTGAATAAGCTAGGTGTTAAGGCAGAAGGCGATGA  
 CAAAGTTGTTATTACTTTATCTAGTCCGTCTCCGCAATTCATCTACTACCTT  
 GCATTCATACTTCATGCCACAAAAACAAGAAGTTGTTGAAAAATATGG  
 AAAAGATTACGCAACTACTTCAAAAAATACAGTTTACTCAGGACCATATA  
 CTGTTGAAGGTTGGAATGGTTCGAATGGTACTTTCACGCTGAAGAAAAAC  
 AAAAATTATTGGGACGCTAAAAATGTAAAAACAAAAGAAGTTTCGCATCCA  
 GACTGTAAAAAACAGATACCGCCGTTCAAATGTATAAACGTGGTGAGT  
 TAGATGCAGCTAATATCTCAAATACTTCTGCTATTTATCAAGCTAATAAAA  
 ATAATAAAGATGTCACAGATGTTCTAGAAGCGACCACTGCCTATATGGAA  
 TATAATACTACTGTTCTGTGAAAGGGCTTGATAATGTTAAGATTTCGTCGC  
 GCCTTAAACTTAGCAACTAACCGTAAAGGAGTTGTTCAAGCAGCCGTTGAT  
 ACAGGCTCAAACCGGCAATTGCTTTTGCACCTACTGGTTTAGCCAAAACA  
 CCAGATGGAAGTATTTGGCAAAATATGTTGCCCCAGGTTATGAATATAAT  
 AAAACTGAAGCAGCAAAACTCTTTAGACTA

MAKQKNNWRRVGVGVLTLASVATLAACGSKSASQDSNGAINWAIPTEINTLD  
 LSKVTDITYSNLAIGNSSSNFLRLDKDGKTRPDLATKVDVSKDGLTYTATLRKG  
 LKWSDBGSKLTAKDFVYSWQRLVDPKTASQYAYLAVEGHVNLADKINEGQEK  
 DLNKLGVKAEGDDKVVITLSSPSPQFIYYLAFTNFMPQKQEVVEKYGKDYAT  
 TSKNTVYSGPYTVEGWNGSNGTFTLLKKNKNYWDANKNVKTKEVRIQTVKKPD  
 TAVQMYKRGELDAANISNTSAIYQANKNNKDVTDVLEATTAYMEYNTTGSV  
 KGLDNVKIRRALNLATNRKGVVQAAVDTGSKPAIAFAPTGLAKTPDGTDLAK  
 YVAPGYEYNKTEAAKLFRL

Sequence description:

- A) Length: 1143 bp - 381 aa (partial sequence)
  - B) Shine-Dalgarno sequence precedes ATG codon.
- Possesses a potential leader peptide sequence.

FIG. 1<sub>CONT'D</sub>

51 / 110

ID-137

Clone 2-c67

TTGAGAGTTTATGAAAATAAAGAAGAGTTGAAAAAAGAAATAAGTAAAAC  
ATTTGAGAAATACATTATGGAATTTAATAA  
TATTCCAGAGAATCTAAAAGATAAAAGAATTGATGAAGTTGATAGAACTC  
CAGCAGAAAACCTTTCTTATCAGGTTGGCT  
GGACCAACTTGGTTCTTAAATGGGAAGAAGATGAAAGAAAGGGACTTCAA  
GTA AAAACACCATCGGATAAATTT

MRVYENKEELKKEISKTFEKYIMEFNNIPENLKDKRIDEVDRTPAENLSYQVG  
WTNLVLKWEEDERKGLQVKTPSDKF

Sequence description

- A) Length: 234 bp - 78 aa (partial sequence)
- B) TTG start codon is preceded by a potential Shine-Dalgarno sequence. No obvious leader peptide sequence.

ID-138

Clone 2-c70

ATGTCAAAGTTTGATAGTCAGAAAATAATTACTCCGATTATGAAGTTTGTC  
AATATGCGAGGGATTATTGCACTCAAAGATGGCATGCTAGCAATTTTACCA  
CTAACAGTTGTTGGGAGTCTCTTTTAAATATTAGGGCAGCTTCCATTT

MSKFDSQKIITPIMKFVNMRGIIALKDGMLAILPLTVVGSFLILGQLPF

Sequence description

- A) Length: 150 bp - 50 aa (partial sequence)
- B) ATG start codon is preceded by a potential Shine-Dalgarno sequence. Possesses a potential

FIG. 1 CONT'D

52 / 110

leader peptide sequence.

ID-139

Clone 2-c71

GAGACCACTTCATCAGTTAAACCAGCAGGAATTGACCGTATCAATCATACC  
TCAACACCCCCGAAGAAAACCTACCCCAACATTGCAACGACGCATAGCTT  
CAAAGATCGTTGTGATACTTTAGAAAGAATTACAAATGAAGACATTGATGT  
TTGTTCTGGATTCAATTTGTGGTATGGGAGAGAGCGATGAGGGGCTCATCAC  
ATTAGCTTTCAGACTAAAAGAACTGAACCCCTATTCTATCCCTGTCAATTTT  
TACTTGCTGTTGAAGGAACACCTCTTGGAAAATATAACTATTTGACTCCC  
ATTAAATGCTTAAAAATTATGGCCATGTTGCGTTTTGTTTTTCCTTTCAAGG  
AATTAAGATTAAGTGCTGGACGGGAGGTCCATTTTGAGAATTTTGAATCAT  
TAGTCACCTTACTTGTTGACTCAACTTTTTTGGGAAATTACCTAACAGAGG  
GGGGTCGCAATCAACATACCGATATTGAATTCTTGGAAAAATTACAATA  
AATCATACTAAAAAGGAATTAATTT

ETTSSVKPAGIDRINHTSTPPKKTTPNIATTHSFKDRCDTLERJHNEDIDVCSGFI  
CGMGESDEGLITLAFRLKELNPYSIPVNFLLAEGTPLGKYNLYLTPIKCLKIMA  
MLRFVFPFKELRLSAGREVFHFNESLVTLVDSTFLGNYLTEGGRNQHTDIEF  
LEKLQLNHTKKELI

Sequence description:

- A] Length: 535 bp - 178 aa (partial sequence)
- B] N- and C-termini require verification

ID-140

Clone 2-c73

ATGCCGGTTTGGACTGCACAGTCTATTCCAAAGGCATTTTGTAGAAAAGCAT  
AATACTAAGGAAGGCACCTGGGCAAACTAACCATTCTAAGTGGTTCTTTA  
GTATTTTACCAGTTATCTCCTGATGGAGAGGAAATCTCGCGGCATATTTT

FIG. 1 CONT'D

53 / 110

GATGCTAGTAGTGATATTCCTTTTGTGATCCACAAGTCTGGCATAAAAGTT  
TCGCCGAATAGTCCAGACTTAAGTTGCTATCTAACTTTTTACTGCCAAAAA  
GAAGATTACTTCCATAAAAAATATGGTCTCACGCGCACACATTCTGAGGTT  
ATCGCCAGTGCACCTCTCTTATCTGAGAAGAGTAATATATTAGACCTTGGG  
TGTGGTCAAGGGCGAACTCACTTTATTTATCGCTGCTGGGACATCAAGTG  
ACTTCTGTCTGATTCAAACGGACAGAGCCTTGTAGCTTTAGAAAATATGGCA  
TTAGAAGAAGAGCTTCCTTACAATATAAAAAGGTATGATATTAATACTACT  
GCTATTGAAGGGCACTATGATTTTATTTTATCAACTGTGGTATTTATGTTTT  
T

MPVWTAQSIPKAFLEKHNTKEGTWAKLTILSGSLVIFYQLSPDGEEISRHFIDAS  
SDIPFVDPQVWHKVSPNSPDLSCYLTFYCQKEDYFHKKYGLTRTHSEVIASAP  
LLSEKSNILDLGCGQGRNSLYLSLLGHQVTSVDSNGQSLVALENMALEEELPY  
NIKRYDINTTAIEGHYDFILSTVVFMF

Sequence description:

- A) Length: 563 bp - 187 aa (partial sequence)
- B) N- and C-termini require verification

ID-141

Clone 2c76

ATGACAAAGCAAATAATTGCCATTTGGGCTGAAGATGAAGACCATTGAT  
TGGAGTTAATGGCGGTTTACCATGGAGGCTTCCTAAAGAGTTACATCACTT  
CAAAGAAACGACCATGGGGCAGGCTTTGCTTATGGGACGAAAGACCTTTG  
ATGGAATGAACCGTCGTGTTTTACCTGGTAGAGAGACAATCATCTTAACAA  
AAGATGAACAATTCCAAGCAGATGGAGTGACAGTCCTAAATAGTGTTGAA  
CAAGTTATAAAATGGTTTCAGGAACATAATAAGACCTTATTTATTGTAGGT  
GGTGCAAGTATTTATAAAGCATTTCTGCCTTATTGTGAAGCAATCATAAAA  
ACTAAAGTTCATGGAAAATTCAAAGGTGATACCTATTTTCCTGATGTTAAT  
CTATCTGAGTTT

MTKQIIAIWAEDDHLIGVNGGLPWRLPKELHHFKETTMGQALLMGRKTFDG  
MNRRLVLPGRETIILTKDEQFQADGVTVLNSVEQVIKWFQEHNKTLFIVGGASI  
YKAFLPYCEAIKTKVHGKFKGDTYFPDVNLSEF

FIG. 1<sub>CONT'D</sub>

54 / 110

Sequence description:

- A] Length: 417 bp - 139 aa (partial sequence)
- B] ATG start codon is preceded by a Shine-Dalgarno sequence. No leader peptide sequence

ID-142

Clone 2-c78

TTGTGGCCAAACTGTGCCCCGCTTATTAATAGCACTTTGTTACCATTTGAA  
GATATCTTAACATCAGGTGCTCATAGCAACCCTATTTTAATGGGGGTTATA  
CTTGGCGGGACAATTGTAGTAGTGGCGACAGCACCACCTTTCTTCTATGGCA  
TTGACAGCTATGCTAGGATTAACCGGAATGCCTATGGCTATAGGAGCCTTG  
TCTGTCTTTGGTTCGTCAATTTATGAATGGTGTACTTTTCCATAAAATTAAC  
TTGGAAGTCGTAAAGATAATATAGCTTTTGCTGTTGAGCCTCTAACTCAAG  
CTGACGTGACTTCAGCTAACCCTATTCCAATCTATGTCATAATTTTGTG  
TGGTGCAGCTTGTGGTATTTTAATTGCCTTGATGAAATTAGTTAATGATACT  
CCTGGAACAGCGACACCAATTGCAGGATTTGCTGTCATGTTTGCCTATAAC  
CCAATGATAAAAGTACTAATAACCGCTCTAGGTTGTATTATCCTATCTTTA  
CTAGCAGGCTATTTTGGAGGCATTGTTTTT

MWPNCAPLINSTLFTIEDILTSGAHSNPILMGVILGGTIVVVATAPLSSMALTA  
MLGLTGMPMAIGALSVFGSSFMNGVLFHKLKLGSRKDNIAFAVEPLTQADVT  
SANPIPIYVTNFBVGGAAACGILIALMKLVNDTPGTATPIAGFAVMFAYNPMIKVL  
ITALGCIILSLLAGYFGGIVF

Sequence description:

- A] Length: 540 bp - 180 aa (partial sequence)
- B] N- and C-termini have yet to be elucidated

ID-143

FIG. 1 CONT'D

55 / 110

Clone 2-c80

ATGTTTTTAAGTATAATGGCAGGTGTCATAGCATTGTCTGACAGTTATT  
GCCATTCCACGCTTCATTAAGTTTTACCAATTGAAGAAAATTGGCGGGCAA  
CAAATGCATGAAGATGTCAAACAACATCTAGCCAAAGCAGGTACGCCGAC  
AATGGGAGGAACGGTATTTT

MFLSIMAGVIAFVLTVIAIPRFIKFYQLKKIGGQQMHEDEVKQHLAKAGTPTMG  
GTVF

Sequence description:

- A) Length: 172 bp - 57 aa (partial sequence)
- B) Shine Dalgarno sequence precedes 'ATG' start codon. Possesses a potential leader peptide sequence.

ID-144

Clone 3-83

ATGAAACCATATTTATCTTTTATTGGTAGAACGTTATTATACTTCGGTATTT  
TATTGTTACTAATTACTTTTTTGCATACCTTGGTCGCGGACAAGGCAGTTT  
TATTATAA

MKPYLSFIGRTLTYFGILLLLIYFFAYLGRGQGSFIY

Sequence description:

- A) Length: 113 bp - 37 aa (partial sequence)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a potential leader peptide sequence.  
This orf is not in frame with nuc

FIG. 1<sub>CONT'D</sub>

56 / 110

ID-145

Clone 3-86

ATGTCATATTTTAGAAATTACTGGTATCGTTTTGGAGCAATTTTATTTATTA  
TTTTAGCAGTAATATTGCTTGTITTTAGACCTGACTGGTCAATGCTTCACTA  
TCTATTGTATTTTACTTTATGGCACTTCTAGCGCATCAATTTGAAGAATAT  
CAGTTTCCCGGTGGGGCATCACCTATCATTAACTATGTTGTTTATGATGAA  
GAAGAGCTGATGGATTGTTTTCCAGGCAATACTCAGTCTATTATGTTGGTT  
AATACTATTGCTTGGTTGCTTTACATTGCTAGTATTGCTTTTCCTCAAGCTT  
ATTGGCTTGGATTAGGAGTCATGTTCTTTAGTCTAACGCAGCTCTTGGGTC  
ATGGTTTTTCAGATGAATATTAACTTAAACTTGGTATAATCCTGGTCTAG  
CAACGACAGTATTTCTCCTAGTACCAATAGCTTGCGCATACATCTATCAAG  
CTAGTGCAGAAGGAATGCTCACTTGGGGAGATTGGCTAGGTGGTTTTATCA  
TGTTGATTGTCTGTGTACTAAGTAGCATTATTGCACCTGTACAGCTATTGAA  
GGATAAGGAGACCAATTATATTATTAGTCCTTGGCAAATGGACCGTTTTCA  
TAAGGTCGTTAATTTTGTAAGGATAAAAAAATAA

MSYFRNYWYRFGAILFIHLAVILLVFRPDWSMLHYLLYFYFMALLAHQFEEYQ  
FPGGASPIINYVYDEEELMDCFPGNTQSIMLVNTIAWLLYIASIAFPQAYWLG  
LGVMFFSLTQLLGHGFQMNILKLTWYNPGLATTVFLVPIACAYIYQASAEG  
MLTWGDWLGGFIMLIVCVLTSHAPVQLLKDKETNYIISPWQMDRFHKVVNFV  
RIKK\*

Sequence description:

A) Length: 651 bp - 219 aa (full length gene)  
B) Putative ATG start codon is preceded by a  
typical Shine-Dalgarno sequence. Possesses a  
potential leader peptide sequence.

ID-146

Clone 3-c88

FIG. 1<sub>CONT'D</sub>

57 / 110

ATGCCACTTACAGCACTTGAAATTAAAGATAAAACATTTTCATCAAAATTT  
CGCGGTTATAGCGAAGAAGAAGTT

MPLTALEIKDKTFSSKFRGYSEEEV

Sequence description:

- A) Length: 75 bp - 25 aa (partial sequence)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No leader peptide

ID-147

Clone 3-90

ATGTCACTTTTTCAAGAAAAAATTGCTTACAATTGCGCTAAAAAGGAAGCG  
CTTTATAAAGAGAGTTTAGGACGCTACGCCTTGAGATCAATGCTAGCAGG  
GGCTTATTTGACAATGAGTACTGCTGCCGGTATCGTCGCAGCTGATACTAT  
TGGTAAAATTTCTCCTGCTCTATCAGGTTTTGTATTTGCTTTCATCTTTAGTT  
TTGGACTTATTTATGTTTTAATATTTAATGGTGAATTGGCGACATCTAATAT  
GCTTTATCTCACTGCAGGAGCCTATAATAAAAAATATCTCTTGGAAAAAAGC  
CATAACAATTTTAATTTATTGTACTTTTTTCAACCTCGTTGGTGCTTGTATA  
TTAGCTTGGTTGTTTAA

MSLFQEKIAYNCAKKEALYKESLGRYALRSMLAGAYLTMSTAAGIVAADTIG  
KISPALSGFVFVAFIFSFLIYVLIFNGELATSNMLYLTAGAYNKNISWKKAITILI  
YCTFFNLVGACILAWLF

Sequence description

- A) Length: 406 bp - 125 aa (partial sequence)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possible leader peptide

FIG. 1<sub>CONT'D</sub>

58 / 110

ID-148

Clone 3-92

AAGTTACAAGCGACTGAAGTTAAGAGCGTTCCGGTAGCACAACCAAGCTTC  
 AACAAACAAATGCAGTAGCTGCACATCCTGAAAATGCAGGGGCTCCAACCTC  
 ATGTTGCAGCTTATAAAGAAAAAGTAGCGTCAACTTATGGAGTTAATGAA  
 TTCAGTACATACCGTGCGGGAGATCCAGGTGATCATGGTAAAGGTTTAGC  
 AGTTGACTTTATTGTAGGTAAAAACCAAGCACTTGGTAATGAAGTTGCACA  
 GTACTCTACACAAAATATGGCAGCAAATAACATTTTCATATGTTATCTGGCA  
 ACAAAGTTTTATTCAAATACAAATAGTATTTATGGACCTGCTAATACTTG  
 GAATGCAATGCCAGATCGTGGTGGCGTTACTGCCAACCCTATGACCACGT  
 TCACGTATCATTAA

KLQATEVKSPVAQPASTTNVAHAHPENAGLQPHVAAYKEKVASTYGVNEF  
 STYRAGDPGDHGKGLAVDFIVGKNQALGNEVAQYSTQNMAANNISYVIWQQ  
 KFYSNTNSIYGPANTWNAMPDRGGVTANHYDHVHVSF

Sequence description

- A] Length: 419 bp - 139 aa (partial sequence)
- B] N- and C-termini have yet to be determined

ID-149

Clone 3-94

ATGATTCCAGTAGTTATTGAACAAACAAGTCGTGGTGAACGTTCTTATGAT  
 ATTTACTCACGTCTTTTAAAAGATCGTATTATTATGTTGACAGGCCAAGTT  
 GAGGATAATATGGCCAATAGTATCATTGCACAGTTATTGTTTCTCGATGCA  
 CAAGATAATACAAAGGATATTTACCTTTATGTCAATACACCAGGTGGTTCA  
 GTATCGGCTGGACTTGCTATTGTGGACACCATGAACTTCATTAAATCGGAC  
 GTACAGACGATTGTTATGGGGATGGCTGCTTCGATGGGAACCATTATTGCT  
 TCAAGTGGTGCTAAAGGAAAACGTTTTATGTTACCGAATGCAGAATATATG

FIG. 1<sub>CONT'D</sub>

59 / 110

ATCCACCAACCAATGGGCGGAACAGGCGGAGGTACACAGCAATCTGATAT  
 GGCTATCGCTGCTGAGCATCTTTTAAAAACGCGTCATACTTTAGAAAAAAT  
 CTTAGCTGATAATTCTGGTCAATCTATTGAAAAAGTCCATGATGATGCAGA  
 GCGTGATCGTTGGATGAGTGCTCAAGAACACTTGATTATGGCTTTATTGAT  
 GCTATTATGGAAAAATAATAATTTACAATAATAGATTTAAAAGAGTTGAGTT  
 TACCAACTCTTTTTTTATTTGTTGGAATTATGTTATAATCTTAGTAATTACA  
 GATATGACGCAGAAAGGAAAAAATTATTGA

MIPVVIEQTSRGERSYDIYSRLLKDRIIMLTGQVEDNMANSIIAQLFLDAQDN  
 TKDIYLYVNTPGGSVSAGLAIVDTMNFIKSDVQTIVMGMAASMGTHASSGAK  
 GKRFLPNAEYMIHQPMGGTGGGTQQSDMAIAAEHLLKTRHTLEKILADNSG  
 QSIEKVHDDAERDRWMSAQEHLIMALLMLLWKIIIYNNRFRKVEFTNSFFICW  
 NYVILVITDMTQKGKNY\*

#### Sequence description

- A) Length: 693 bp - 231 aa (full length gene)
- B) Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No leader peptide. Significantly, it would appear to have a very hydrophobic C-terminus.

ID-150

Clone 2-c86

ATGAAACCAAAAATTATTGGTGTACTTGGTCTAGGAATATTTGGACAAACA  
 CTCGCACAAGAACTAAGTAACCTTTGAACAAGATGTTATTGCTATTGACAGC  
 AATCCTGAAAATGTACAAGCTGTCGCCGAAGT  
 TGTTACAAAAGCAGCTATCGGAGACATTACTGATTTAGCTTTCCTAAAACA  
 CATCGGGATCAGTGACTGTGATACTGTTATTATTGCTACAGGAAACAGTTT  
 AGAGAGCTCAGTATTGGCCGTAATGCACTGTAAAAAGTTAGGCGTCCCAC  
 AAGTTATTGCTAAAGCTCGAAACCTTGTATACGAAGAAGTACTTTATGAAA  
 TTGGTGCTGATTTGGTTATCTCTCCGGAGCGAGAATCTGGGCAAAATGTTG  
 CTGCAAACCTCATGAGAAATAAAATTACAGATGTCTTCCAGATTGAATCTG  
 ATATTTCTGTCATTGAATTT

MKPKIIGVLGLGIFGQTLAQELSNFEQDVIAIDSNPENVAEAVVTKAAIGDI  
 TDLAFLKHIGISDCDVTIIATGNSLE

FIG. 1<sub>CONTD</sub>

60 / 110

SSVLAVMHCKKLGVPQVIKARNLVYEEVLYEIGADLVISPERESGQNVAAAN  
LMRNKITDVFQIESDISVIEF

Sequence description:

A] Length: 459 bp - 153 aa (partial sequence)  
B] Putative ATG start codon is preceded by a  
typical Shine-Dalgarno sequence. Possesses a  
potential leader peptide sequence.  
This orf is not in frame with nuc

ID-151

Clone 2-c88

GTGCGTTATAGTAAAGAGATTATTCAGTTAGCTATACCAGCTATGATTGAA  
AATATCTTACAAATGCTCATGGGAGTAGTTGATAATTATCTAGTGGCTCAG  
TTAGGTGTTGTAGCAGTATCAGGTGTTTCAGTTGCTAATAATATAATTACT  
ATTTATCAAGCTATTTTTATAGCTTTAGGGGCGAGTATAGCAAGTCTATTG  
GCCAAGTCGTTAGCAGGTAGTGAGAAGGATGATGCAATTTTCAGTATGTTCT  
CAAGCCATTTTTCTAACATCACTGATAGGGGCGAGTATTAGGAATTATCTCG  
ATTGTTTTTGGACAAACTTTCTTT

MRYSKEIIQLAIPAMIENILQMLMGVVDNYLVAQLGVVAVSGVSVANNIITIY  
QAIFIALGASIASLLAKSLAGSEKDDAISVCSQAIFLTSLIGAVLGHISIVFGQTFF

Sequence description

A] Length: 330 bp - 110 aa (partial sequence)  
B] Putative GTG start codon is preceded by a  
typical Shine-Dalgarno sequence. May have a  
leader peptide

ID-152

FIG. 1<sub>CONT'D</sub>

61 / 110

Clone 2-c92

TTGATTAACAAGTATTCGTGCTTTTTGAAGAGGATTCTCCATAATAATACT  
CCTTTAATAGTTATCGTGAGAAGTATTTTAAAGAAAAACCGCCAAGGTAG  
AGCGACATTTCTGCCTTTAACTACAATAAAACCAAGAGAATTAGCACAAAC  
ATTATCTCTCAAATTACAAAGTTCTCAAGGGTTTTTAGGAATAGCTAGTG  
AATTGGTAACCTATGATCAACGCTTGTCAAACATTTTT

MINKYSCFLKRILHNNTPLIVIVRSILKKNRQGRATFLPLTTIKPRELAQHYLSK  
LQSSQGFLGIASELVTYDQRLSNIF

Sequence description

- A] Length: 240 bp - 80 aa (partial sequence)
- B] No obvious Shine Dalgarno sequence precedes the Putative TTG start codon

ID-153

Clone 2-c94

TTGTTGACTCACAAAAATATATTATTAACCATTATATTTGGATTATTTATGA  
TTATATTATCAGCATGTGGTATGTCTAATAAGGAAATGGCTGGTATTGATA  
ATTGGGAACATTATCAAAAGGAAAAGAAAATTACTATTGGATTGATAAT  
ACTTTTGTTCCCTATGGGATTTGAAAGTCGTTCTGGTGACTATACCGGCTTTG  
ATATTGATTTAGCTAATGCTGTTTTTAAAGAATACGGTATTTTCAGTGAAAT  
GGCAGCCTATTAAGTGGGATATGAAAGAACTGAACTTAATAATGGTAAT  
ATAGACCTTATTTGGAATGGTTATTCAAAAACGGCAGAACGTGCTAAAAA  
AGTCGCTTTTACAAACCCATATATGAATAATCATCAAGTAATTGTTACTAA  
AACTTCATCACATATTAATAGTATTAAGGATATGAAGGGGAAAAAACTAG  
GAGCCCAGTCGGGTTTCATCTGGTTTTGATGCTTTTAAACGCTAAACCTGATA  
TTTTAAAAAAGTTTGTAAAAGGAAAAGAAGCAGTTCAATACGATACTTTC  
ACTCAGGCTTTGATTGATTTAAAAAATAACCGTATTGATGGTCTTTTGATT  
GATGAAGTTTATGCTAACTATTATTAAAGCAAGAAGGAA

FIG. 1<sub>CONT'D</sub>

62 / 110

MLTHKNILLTIIFGLFMIILSACGMSNKEMAGIDNWEHYQKEKKITIGFDNTFV  
PMGFESRSGDYTGFDIDLANAVFKEYGISVKWQPINWDMKETELNNGNIDLI  
WNGYSKTAERAKKVAFTNPYMNNHQVIVTKTSSHINSIKDMKGKKLGAQSG  
SSGFDAFNAKPDILKKFVKGKEAVQYDTFTQALIDLKNNRIDGLLIDEVYANY  
YLKQEG

Sequence description

A] Length: 649 bp - 216 aa (partial sequence)  
B] TTG start codon is preceded by a possible  
typical Shine-Dalgarno sequence. Has a  
leader peptide

ID-154

Clone 2-c100

ATGAAAATTTGGAAAAAATAACCTTAATGTTTTCTGCAATTATTTTAACA  
ACAGTAATTGCATTGGGAGTCTATGTTGCCTCAGCTTATAATTTTTCGACTA  
ATGAATTGTCTAAGACTTTT

MKIWKKITLMFSAIILTTVIALGVYVASAYNFSTNELSKTF

Sequence description

A] Length: 123 bp - 41 aa (partial sequence)  
B] ATG start codon is preceded by a potential  
typical Shine-Dalgarno sequence. Has a  
typical leader peptide

ID-155

Clone 2-c1

FIG. 1<sub>CONT'D</sub>

63 / 110

ATGAAAAAACAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA  
 ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA  
 GGAATACCAGGCAGAACAGAAATTTAAGTCATACTTTAAATATATATCAG  
 ATAAAAATAACTATTTAGATAATATAAAAGTTTATTACTTTTCTATAAGTA  
 TTTCTAAAGATGTACAAGATAAAGTCAGTGAAACAACAACCTTGTTTCATATA  
 GACTAGAAAAGCAAAAGAATCAAGAGTTCATTGGTAATTTTGAACATGAA  
 GTTAGTGAATCTAGTCAATATTCAACCGAAGTTAAAAATCAAATACAGTAT  
 CCAATCCAGTATAAAGATAATTCAATTCGTTTTACTGAAAAAACACCGTCA  
 GAACGTTATGATGAGTTTGTGTTTTAGTTCATTTGATTCTTCATTATTA  
 AATATAAAATATATGATTACTTACTAAAACATCCCGAAACTGAATTA  
 GGTGTTTCCTATAAGATTCTCTATAAATTCTGAAATTGTAGCCCCCTTTATAA  
 ATCAATTAAATATAAAAAATCCTAAAAAATCATCTATTTCCGTTACAAAAA  
 CGGAAAGTAAAGAATATTATTATACAATCAGTATTGATACTGATTCTGAGA  
 TATATTCTATATTCTGAAGGTATTCAT

MKKQRLLLLFGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYFKYISDKN  
 NYLDNIKVYYFSSISISKDVQDKVSETTTCSYRLEKQKNQEFIGNFEHEVSESSQ  
 YSTEVKNQIQYPIQYKDNSIRFTEKTPSERYDEFVFSFDSSLLKKYKIYDYLLK  
 HPETELKGVSYPINSEIVAPFINQLNIKNPKKSSISVTKTESKEYYYTISIDTDS  
 EIYSIFEGIH

#### Sequence description

- A] Length: 687 bp - 229 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Has a typical leader peptide. C-terminus has yet to be verified

ID-156

Clone 2-c5

ATGACATTTGACACCATTGATCAATTAGCGGTTAATACAGTCCGCACGCTT  
 TCTATTGATGCTATCCAAGCAGCAAATTCTGGGCACCCAGGTCTTCCTATG  
 GGAGCTGCGCCTATGGCTTATGTGCTTTGGAATAAATTCTTAAATGTAAAC  
 CCAAAAACAAGTCGCAATTGGACAAACCGTGACCGTTTTGTACTTTCAGCT

FIG. 1 CONT'D

64 / 110

GGGCATGGTTCAGCTCTTCTTTATAGCCTACTTCATTTAGCTGGCTATGATT  
TATCAATTGATGATT

MTFDTIDQLAVNTVRTLSDAIQAANSGLPLMPGAAPMAYVLWNKFLNVNP  
KTSRNWTNRDRFVLSAGHGSALLYSLHLAGYDLSIDD

Sequence description

- A] Length: 272 bp - 90 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. No obvious leader peptide

ID-157

Clone 2-c8

ATGAGAACACTATTTAGAATGATATTTGCTATTCCAAAGTTTATCTTTAGA  
TTGATTGGGAATATCATTTGGGGAATATTCAAGACAGTTCTTGTTATTGCG  
ATTATTTTATTTGGCTTGTATTACTATGCGAATCACAGTCAATCAGAATTTG  
CTAATCAACTTAGTGACATTATTCAGACAGGAAAAACATTTTT

MRTLFRMIFAIPKFIFRLIWNIIWGIFKTVLVIAIILFGLYYYANHSQSEFANQLS  
DIIQTGKTF

Sequence description

- A] Length: 197 bp - 65 aa (partial sequence)
- B] ATG start codon is preceded by a potential typical Shine-Dalgarno sequence. Possesses a leader peptide

ID-158

FIG. 1<sub>CONT'D</sub>

65 / 110

Clone 2-c9

ATGTC AAAAAAATAATATTAGGAATTTTATCTCTTTTATCTGTCGTTACTT  
TGGTGGCGTGTGGTTCATCAGACAAACAGCTACAAGATAAAGTTGAGAAA  
AAAGGGAAGTTAGTTTTAGCGGTGAGTCCAGATTATGCTCCCTTTGAGTTT  
MSKKIILGILSLLSVVTLVACGSSDKQLQDKVEKKGKLVLA VSPDYAPFEF

Sequence description

A] Length: 153 bp - 51 aa (partial sequence)  
B] ATG start codon is preceded by a potential  
typical Shine-Dalgarno sequence. Possesses a  
leader peptide (not in frame with nuc)

ID-159

Clone 2-c10

ATGAAAAATCAAAGACTATTACTGCTTTTTTGGAGGCTTATTAATAATGATA  
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA  
GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTT  
MKNQRLLLLFGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYF

Sequence description

A] Length: 139 bp - 46 aa (partial sequence)  
B] ATG start codon is preceded by a potential  
typical Shine-Dalgarno sequence. Possesses a  
leader peptide

FIG. 1 CONT'D



FIG. 1 CONTD

68 / 110

examination.

ID-163

Clone 2-c25

TTGAAAAAATCCAAGAGAAGCCGTAAGGCAGTGACAACAAGTGGTGAGA  
AGACTTTACTTGAGGATTTGGCAAAAATGAATTCCTAGACGAAGTCATTA  
ATGTTATGGTTTTATATACCTTGAATAAGACAAAATCTGCTAACTTAAATA  
AGGCCTATATCATGAAAGTTGCTAATGATTTTGCCTTTCAGAATGTTATGA  
CGGCCGAAGATGCTGTGCTTAAAATTCGTGATTTTTCAGATCAAAAAGTAA  
GGACTAAACAGAAACGAAGAAGAAACAATCGAATGTTCTGAATGGAGT  
AATCCTGATTATAAAGATGAGGTTAGCCCAGAAAAAGAAATTGAATTAGA  
ACAGTTT

MKSKRSRKA VTTSGEKTLLDLAKMNFLEVINVMVLYTLNKTKSANLNK  
AYIMKVANDFAFQNVMTAEDAVLKIRDFSDQKVRTKTETKKKQSNVPEWSN  
PDYKDEVSPKEIELEQF

Sequence description

- A) Length:360 bp - 120 aa (partial sequence)
- B) N- and C- termini require verification.

ID-164

Clone 2-c28

ATGACGAATCATATTACTAAACTGATAGAAAATAGCGGAAAAAAATTGAC  
AGAAATTAGCGAAGCTACAGATATAGCCTATCCTACACTTTCTGGATACAA  
TCAAGGAATCCGCAAACCTAAAAAAGATAATGCTGAAAAATTGGCAAAAT  
ACTTTAATGTTTCCGTCGCTTACATTATGGGACTTGATAGCAACCCACATG  
CTCCATCAAATCTT

MTNHITKLIENSGKKLTEISEATDIA YPTLSGYNQGIRKPKKDNAEKLAKYFNV  
SVAYIMGLDSNPHAPSNL

FIG. 1<sub>CONT'D</sub>

69 / 110

# Sequence description

- A) Length:218 bp - 72 aa (partial sequence)
- B) ATG start codon is preceded by an obvious Shine Dalgarno sequence. No obvious leader peptide.

ID-165

Clone 2-c29

TTGATGAAAAGGAATAAACATTTACCGTTAACAGAACTACCTATTATATT  
TTATTAGCTTTGTTTGAGGAAGCGCATGGCTATGCTATTATGAAAAAAGTT  
GAAGAAATGAGTGGCGGTGATGTTAGAATAGCCGCAGGGACAATGTACGG  
TGCCATTGAAAATTTACTTAAACAAAAATGGATAAAGTCTATCTCAAGTGA  
CGATAGAAGAAGAAAAGTTTATATTATTACTGAGACAGGAAAAGAAATAG  
TAGAACTTGAAACGAATCGATTAAAGAAAGTTACTTAATACTGCTAATCAGT  
TGGGTTTTGGAGGAGATGGTTATGATAAAGTTT

MMKRNKHLPLTETTTYILLALFEEAHGYAIMKKVEEMSGGDVRIAAGTMYG  
AIENLLKQKWIKSISSDDRRRKVYIITETGKEIVELETNRLRKLLNTANQLGFG  
GDGYDKV

# Sequence description

- A) Length:337 bp - 112 aa (partial sequence)
- B) TTG start codon is preceded by an obvious Shine Dalgarno sequence. Actual start codon may ATG that comes immediately after the TTG. Potential leader peptide.

ID-166

FIG. 1 CONT'D

70 / 110

Clone 2-c35

CCCATTACTGGTGAGTTAATAGCTGAGAAATTAGGAGTACCAAGAGCAGC  
 ACTAAGGTCTGATTTGCGGGTTTTAAGTATGCTAGGTATCATAGATGCAAA  
 ACCTAAGGTTGGTTATTTTTATTTAGGACAGTATCATGCTTCAATAGGGAC  
 AAGTCATTTTGAAAAGATGACAGTTTCAGAAATTATGGGGATCCTTCTGAC  
 AGTTCATCAAAAAGATTCAGTTTATGATGTTATTGTACATATTTTTATGGA  
 AGATGCTGGTTGTGCTTTTATCTTGGATGATGATGATTTTCTCTGTGGAGTC  
 GTGTCACGTAAAGATTTACTAAAAACCAGTATTGGCGGAGGAGATCTTTCT  
 AAAATGCCAATAGGAATGGTGATGACACGTATGCCACACGTGACAACTGT  
 TTTAGAAAATGAAAGTCTTTTTGCGGCAGCTGATAAATTAGTGAGCAGAA  
 AAGTGGATAGTCTCCCTGTCGTCGTCATGATAAGCAATATCCCGAAAAAT  
 TTA

PITGELIAEKLGVPRALRSDLRVLSMLGIIDAKPKVGYFYLGQYHASIGTSHF  
 EKMTVSEIMGILLTVHQKDSVYDVIVHIFMEDAGCAFILDDDDFLCGVVSRKD  
 LLKTSIGGGDLKMPIGMVMTRMPHVTTVLENESLFAAADKLVSARKVDSL PV  
 VRHDKQYPEKF

Sequence description

- A) Length:511 bp - 170 aa (partial sequence)
- B) N- and C-termini to be determined

ID-167

Clone 2-44

TTGGAAGTCATCATGCAATTTATTTATAGTATTATTGGTATTTTATTGGTAT  
 TAGGAATTGTGTATGCAATTTCTTTCAATCGTAAGAGTGTTTCTCTAAGTTT  
 AATTGGAAAAGCTCTTATCGTTCAATTCATTATTGCGCTAATCTTAGTACGT  
 ATCCCACTAGGCCAACAAAGTTGTTAGTGTTGTTTCAACTGGAGTTACTAAA  
 GTAATCAACTGTGGTCAAGCTGGTTT

MEVIMQFIYSHIGILLVLGIVYAISFNRSVSLSLIGKALIVQFIILILVRIPLGQQ  
 VVSVVSTGVTKVINCGQAG

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

72 / 110

TGCAGCTCTTACTGCTCTCTTTTTAGGATATAAATAGCGTTACTGCGGATACA  
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTAACTGAGGA  
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA  
ATATACCTTTT

MKCIINNINKIKMIIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY  
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

#### Sequence description

- A) Length:264 bp - 88 aa (partial sequence)
- B) There is a Shine-Dalgarno sequence upstream of this sequence. Potential leader peptide sequence

ID-169

Clone 2-47

ATGAAATGTATAATAAATAATATAAATAAAATAAAAAATGATAATTGAGAT  
TTATCATAGAAGGAAACTATTTTGAAATTAAATAAAATCATATTATCTAC  
TGCAGCTCTTACTGCTCTCTTTTTAGGATATAAATAGCGTTACTGCGGATACA  
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTAACTGAGGA  
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA  
ATATACCTTTT

MKCIINNINKIKMIIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY  
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

#### Sequence description

- A) Length:264 bp - 88 aa (partial sequence)
- B) There is a Shine-Dalgarno sequence upstream of this sequence. Potential leader peptide sequence

FIG. 1<sup>CONT'D</sup>

### Clone RS-58b

MGDY YGKKYFG EAAKKDVEHMAKKIINVYKTRLKNNTWLSENTKAMAIAKK  
LDNMRLMIGYPDYPDL YRQYQFDSKASFFENNDNYRKLSNKKTFEENQSNQ  
REHWQMSANAVNAYNDPNTNSIVFPAAIFQSPLYDKTKTVSQNYGAIGAIIGH  
EISHSFDINGMKYDEKGNLHDWWTKEDLNHYKKSTQAMIDQWDGLKADGG  
KVDGKLT LAENIADNGGVMASLEALKTEKIQTIKNFLNHGQVFGVKKQPKNK  
VSPQFSQMFMHHMN\*

**A] Length: 819 bp - 272 aa (full length gene)  
(107 bp of additional DNA sequence (> onwards) is  
also included. While not in-frame with the  
described orf, it also shares strong homology  
with the neutral peptidases.**

FIG. 1 CONTD

74 / 110

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-89 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-89 gene sequence. ID-89 and ID-170 together show homology over their combined entire length with the neutral endopeptidases from *Lactococcus* and *Lactobacillus*. Possesses TTG (possible ATG start codon located 13 bp further downstream) start codon with no obvious signal peptide. Shine Dalgarno sequence not immediately obvious. Possibly located further downstream

ID-171

Clone 2-18/22b (Mod2)

ATGACCATGATTACGCCAAGCTTCATTAAGGTATCTCTAGATGAAACAAAT  
CGTATGATGCGTATGATATCAGATTTATTAAGTTTATCGCGCATTGATAAT  
GAAGTAACGCATTTAGATGTTGAAATGACGAATTTTACAGCTTTCATGACC  
TCAATTTTGAATCGATTTGATCAGATTAGAAATCAAAAAACAGTCACAGG  
AAAAGTTTATGAAATTGTCAGAGATTATCCTCTTAAGTCAATTTGGGTGGA  
AATTGATACAGATAAGATGACTCAAGTGATTGATAACATTTTAAATAATGC  
AGTCAAGTATTCACCAGATGGTGGTAAGATTACAGTTAATCTACGCACAAC  
TAAACGCAGATGATTTTATCAATATCAGACCAAGGCTTAGGTATTCCCAA  
AAAAGATTTACCTCTCATTTTTGATCGTTTTTATCGTGTTGATAAGGCGAGA  
AGTCGTCAACAGGGTGGGACTGGACTTGGTTTGTCAATTGCAAAGAAAT  
TGTTAAGCAGCATAAGGGATTTATTTGGGCTAAGAGTGAGTATGGTAAAG  
GGTCTACTTTTACAATCGTCTTGCCTTATGATAAAGATGCTGTAACCTTATGA  
AGAATGGGAGGACGTTGAAGATTAA

MTMITPSFIKVSLEDETNRMMRMISDLLSLSRIDNEVTHLDVEMTNFTAFMTSIL  
NRFDQIRNQKTVTGKVEIVRDYPLKSIWVEIDTDKMTQVIDNILNNAVKYSP  
DGGKITVNLRTTKTQMILSISDQGLGIPKKDLPLIFDRFYRVDKARSRQQGGTG  
LGLSIAKEIVKQHKGFWAKSEYKGSTFTIVLPYDKDAVTYEEWEDVED\*

Sequence description:

- A] Length: 613 bp - 212 aa (full-length gene possibly)
- B] Possible Shine Dalgarno sequence present upstream of a ATG start codon. May not have yet determined the N- portion of this gene. No obvious signal peptide.

FIG. 1<sub>CONT'D</sub>

75 / 110

ID-172

Clone 2-54balternate (107b)

TTGAAAAAATTATTACTTCTATTCTATTACTTAGTTGCATTTTTTTTATGC  
CAACCATCTCTGCTGAATCTTTTAATGCTTCCGCTAAACATGCCTTAGCAGT  
TGATTTAGATTTCAGGAAAAATCTTGTATGAAAAAGATGCTAACAAACCCG  
CTGCTATTGCTTCCTTGACTAAAATAATGACCGTTTATATGGTCTATAAAG  
AAATTGATAACGGTAACCTCAAGTGAATACCAAAGTAAATATATCTGAC  
TACCCTTATCAACTAACACGCGAATCTGATGCTAGTAATGTTCTTTAGAA  
AAAAGGCGCTATACTGTAAACAACCTCGTGGACGCTGCCATGATTTCTAGT  
GCTAACAGTGCAGCCATTGCTTTAGCTGAACATATTTTCAGGAACTGAAAGT  
AAATTTGTTGATAAAATGACTGCTCAATTGGAAAAGTGGGGAATTCATGAT  
AGCCACCTAGTCAATGCTTCTGGCTTAAATAATAGTATGTTAGGCAATCAC  
ATTTATCCAAAATCGTCACAAAACGACGAAAATAAAATGAGTGCACGTGA  
TATTGCTATTGCTGCCTACCATTTGGTCAACGAATATCCTTCCATTCTTAAG  
ATTACTAGTAAGTCCGTTGCTAAATTTGATAAAGATATTATGCATTCTTAT  
AACTACATGCTACCAGATATGCCTGTCTTTAGACCAGGTATTACAGGTTTG  
AAAAGTGGGACAACGGAATTAGCTGGCCAATCTTTATTGCTACATCTACT  
GAAAGTGGAAATGAGACTACTCACTGTTATTATGCATGCTGATAAGGCCGAT  
AAAGACAAATATGCTCGCTTTACAGCAACTAACTCTCTCTTGAACATATATC  
ACAAACACCTACGAACCTAACCTTGTATTAGCTAAAGGAGCTGCATATAA  
AGGTAAAGAAGCAAGTGTGAGAGACGGAAAAGAACAATCGGTCATCGCT  
GTTGCTAAAAACGATTTGAAAGTAGTACAGAAGAAAAATATCACTAAACA  
AAATCAGTTAAAAATTAACCTTTAAAAAAGAGCTTACTGCTCCTATTACAAA  
AAAAGAGAACCTAGGGAAAGCTTATTACGTTGACCTTAATAAGGTTGGAA  
AAGGCTATCTCATAAAGGAACCTAGCGTTCATTTAGTGGCAAAGATAGT  
ATTGAGCGCAGTTTCTTCCTCAAAGTGTGGTGAATCATTTTGTGCGCTAC  
GTTAACGAAAAACTTTAA

MKKIITSILLSCIFFMPTISAESFNASAKHALAVDLDSGKILYEKDANKPAAIA  
SLTKIMTVYMVYKEIDNGNLKWNTKVNISDYPYQLTRESASNVPLEKRRYT  
VKQLVDAAMISSANSAAIALAEHISGTESKFVDKMTAQLEKWGIHDSHLVNA  
SGLNNSMLGNHIYPKSSQNDENKMSARDIALAAHYHLVNEYPSILKITSKSVAKF  
DKDIMHSYNYMLPDMPVFRPGITGLKTGTTELAGQSFIATSTESGMRLLTVM  
HADKADKDKYARFTATNSLLNYITNTYEPNLVLAKGAAKYKGKEASVRDGKE  
QSVIAVAKNDLKVVQKKNITKQNQLKINFKKELTAPITKKENLGKAYYVDLN  
KVGKGYLEKEPSVHLVAKDSIERSFFLKVWWNHFVRYVNEKL\*

FIG. 1<sub>CONT'D</sub>

76 / 110

## Sequence description:

- A) Length: 1236 bp - 412 aa (full-length gene sequence possibly)  
 B) A possible Shine-Dalgarno sequence precedes the putative 'TTG' start codon. (needs further cloning and sequencing to verify N-terminus)

ID-173

Clone 3-60b

ATGACGCTTCGAGAATTAACAATAGAAGAATTTAAAGAACATTCAGGAAA  
 TTATGATTCACAATCATTTTTACAAACACCTGAGATGGCTAAACTTTTAGA  
 AAAACGCGGCTATGATGTTAGGTATTTGGGATATCAAGTAGAAAATAAAC  
 TAGAGATAATCAGTTTATCTTATATTATGCCAGTCACTGGTGGTTTTCAAAT  
 GAAAATTGATTCAGGACCAGTTCATTCAAATTCTAAGTATCTAAAACAATT  
 TTATAAAGCATTGCAAGGCTATGCCAAATCCAACGGTGTCTAGAATTAAT  
 AGTTGAGCCTTTTGATGATTACCAATTATTCAGTTCGGGAGTTCCTAGT  
 AATCAGGGAAATGATAATCTGATTGAAGATTTTACCAGTTCAGGTTATCAC  
 CATGATGGTTTAACAACCTGGTTTTACTGGTAAATATTTATCTTGGCACTATG  
 TTAAAAATTTAGAAGGTGTCACTTCTGAAACGTTACTATCTTCATTCTCTAA  
 GACAGGACGAGCTTTGGTTAAGAAAGCAATGTCTTTTGGGAATCAAGGTTT  
 GCGTTCCTTAAACGTGATGAGCTACATTTATTTAAAGAGATAACAACCTTCTA  
 CGTCAAATAGACGTGATTATATGGATAAGTCCTTAGATTATTATCAAGATT  
 TTTACGATAGCTTTGAAGGCAAGGCTGAATTTGTGATTGCCACTTTAAATT  
 TTAGAGAATACGACCATAACTTGCAAATAAAAGCTGAAGCATTGGAAAAAT  
 AAGCTT

MTLREL TIEEFKEHSGNYDSQSFLQTPEMAKLLEKRGYDVRYLGYQVENKLEI  
 ISLSYIMPVTGGFQMKIDSGPVHSNSKYLKQFYKALQGYAKSNGVLELIVEPF  
 DDYQLFTSSGVPSNQGNLIEDFTSSGYHHDGLTTGFTGKYLSWHYVKNLE  
 GVTSETLLSSFSTGRALVKKAMSFGIKVRVLKRDELHLFKEITTSTSNRRDY  
 MDKSLDYYQDFYDSFEGKAEFVIATLNFREYDHNLQIKAEALENKL

## Sequence description

- A) Length: 771 bp - 257 aa (partial gene sequence)  
 B) This gene sequence was not identified using the LEEP system. It was identified immediately downstream of the ID-65 gene which was identified by

FIG. 1 CONT'D

77 / 110

LEEP, during cloning and sequence analysis of the full-length ID-65 gene sequence. Sequence Characteristics:  
 No obvious leader peptide sequence  
 Orf is preceded by a potential Shine-Dalgarno sequence.

ID-174

Clone 2-17b (ID-80b)

TTGTCATTAAGTTTGGTTGCAGTGTTAAATCTTATCCCTCCTAAAATCATGG  
 GATCAGTTATTGATGCTATTACAACCTGGAAAATTAACAAGACCACAATTAC  
 TATGGAATTTATTAGGTTTGGTTTTGTCAGCTTTAGCTATGTATGGGCTGCG  
 TTATATTTGGCGTATGTATATTTAGGGACTTCTTACAAATTAGGCCAAGTT  
 GTCAGATACCGTTTATTTGAACATTTACAAAAATGTCTCCTTCTTTTTATC  
 AGAAATATCGTACAGGTGATTTAATGGCGCACGCGACCAACGACATCAAT  
 TCTCTAACACGTCTTGCAGGAGGAGGAGTTATGTCAGCAGTGGATGCCTCT  
 ATCACAGCATTAGTAACGCTTATCACCATGTTCTTTACTATTTTCGTGGCAA  
 ATGACATTAATTGCGGTTATCCCTTTGCCCTTAATGGCCTTAGCACTAGTA  
 AATTGGGGCGAAAAACCCATGAAACCTTCAAAGAATCTCAGGCAGCCCTT  
 TTCAGAATTAATAATAAAGTG

MSLSLVAVLNLIPPKIMGSVIDAITTGKLTTPQLLWNLGLVLSALAMYGLRYI  
 WRMYILGTSYKLGQVVRYRLFHFTKMSPSFYQKYRTGDLMAHATNDINSLT  
 RLAGGGVMSAVDASITALVTLITMFFTISWQMTLIAVIPLPLMALALVNWGEK  
 PMKPSKNLRQPFSELNNKV

#### Sequence description

A) Length: 534 bp - 178 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-80 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-80 gene sequence.

Sequence Characteristics:

No obvious leader peptide sequence

Orf is preceded by a potential Shine-Dalgarno sequence.

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

79 / 110

AATTTGGACTGTATAACATTTTGGGGATGAATAAACGTCAAGTTGCGCGTG  
TAGCTAGTCTAGAGCTGTTTATTATTTATATATTTCTTATTTCTATAGGAAG  
TCTGTTTAGTGCTTTTTTTGCTAAATTTATTTATTTAATTTTGTCAACATTA  
TTAACTATCATGCACTAAATCTTAGTTTAAGTTTATGGCCATTTATTATTTG  
TATCGTTATATTTACAGGTATTTTCTGACTTTAGAAGTTCCAGTTATTCGA  
CATGTTCAATTTATCATCCCCATTAAGTCTTTTTAGAAAGAAACAACAGGGA  
GAAAAAGAACCAAAAAGGTAATCTTATACTTGCAATTTTAGCGTTAGTAGCT  
ATCGCCATCGCTTATACAATGGCTCTTACTTCAGGTAAAGCACCTGCATTA  
GCTGTTATCTATCGTTTCTTCTTGCAGTACTTTTAGTAATTGCTGGTACTT  
ATCTTTTTTATATTAGTTTTATGACATGGTACTTAAAAAGGTTGCGTCAAAA  
CAAGCATTATTATTATAAATCTGAGCATTTTGTATCAACTTCGCAAATGAT  
TTTTCGAATGAAGCAAAATGCAGTAGGGTTAGCAAGTATCACTTTATTAGC  
TGTTATGGCTCTAGTTACTATTGCTACAACAGTCTCACTCTATTCAAATACA  
CAAAATGTTGTTACCGGACTATTTCCAAAATCAGTAAGTTTATCAATAGAT  
AATTCAAAAGGTGACGCGAAAAATATATTTGAAGAAAAGATTTTGAAGAA  
ACTAGGTAAGTCATCTAAGGAAGCTATCACTTATAATCAGACAATGATTTC  
GATGCCAGTTAGTCAATCAAGTGACTTAATATCACATCTA

MNNMFYLKIAWHNLKHSIDQYIPFLLASLLLYSLTCSTLLILMSAVGRDMGTA  
ATVLF LGVIVLSIFAVVMEHYSYNILMKQRSSEFGLYNILGMNKRQVARVASL  
ELFIIYIFLISIGSLFSAFFAKFIYLIFVNIINYHALNLSLSLWPFHICIVIFTGIFLTLE  
VPVIRHVLSSPLSLFRKKQQGEKEPKGNLILAILALVAIAIAYTMALTS GKAP  
ALAVIYRFFFAVLLVIAGTYLFYISFMTWYLRRLRQNKHYYYKSEHFVSTSQM  
IFRMKQNAVGLASITLLAVMALVTIATTVSLYSNTQNVVTGLFPKSVSLSIDNS  
KGDAKNIFEEKILKKLGKSSKEAITYNQTMISMPVSQSDDLISHL

Sequence description:

- A] Length: 1119 bp - 373 aa (partial gene sequence)
- B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-104 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-104 gene sequence. Possible Shine Dalgarno sequence present upstream of a GTG start codon. Possesses a potential leader peptide sequence

ID-177

FIG. 1<sub>CONT'D</sub>

80 / 110

Clone 2-5b (ID-112b)

ATGGTTGAGCCAATTATTTCAATACAAGGACTTCATAAAAAGTTTTGGGAAA  
AATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGT  
GGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAAT  
GAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTG  
ATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGC  
ATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAAT  
ATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAG  
ACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA  
TGCTTATCCAGCAAGCTTATCTGGAGGACAACAACAACGGATTGCTATTGC  
AAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTGATGAACCTACTTCA  
GCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTA  
GCTAAATCTGGTATGACGATGGTTATTGTCATCATGAAATGGGTTTTGCA  
CGTGAAGTAGCGGATCGTGTCATTATTATGGATGCAGGGATTATTGTTGAG  
CAAGGGACCCCTAAGAAAGTATTTGAGCAGACAAAAGAAATCCGCACAAG  
AGACTTCTTAAGTAAAGTATTATAA

MVEPIISIQGLHKSFGKNEVLKGIDLDIHQGEVVVIIGPSGSGKSTFLRTMNLLE  
VPTKGTVTFEGIDITDKKNDIFKMREKMGVMVFQQFNLFNMTVLENITLSPIKT  
KGLSKLDAQTKAYELLEKVGLKEKANAYPASLSGGQQORIAIARGLAMNPDV  
LLFDEPTSAIDPEMVGEVLTVMQDLAKSGMTMVIVTHEMGFAREVADRVIF  
MDAGIIVEQGTPKKVFEQTKEIRTRDFLSKVL\*

Sequence description:

- A] Length: 735 bp - 244 aa (full length gene)
- B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence precedes the 'ATG' start codon. No obvious leader peptide

ID-178

Clone 2-5c (ID-112c)

FIG. 1<sub>CONT'D</sub>

81 / 110

ATGTCTCA<sup>s</sup>TATCAAGAGTGGTTAGAAAACGACTCACTCGGTAAAGATATT  
 AAGTCAGATTTAGAAGCTATTAAAGGAGATGAATCTGAAATTCAGGATCG  
 TTTTACAAAACATTAGAATTTGGAACGGCGGGATTGAGAGGTAAACTTG  
 GAGCAGGAACCAATCGTATGAATACTTATATGGTGGGGAAAGCAGCACAA  
 GCATTAGCTAATCGATTATTGATCATGGCCCTGAAGCTATTGCACGTGGAA  
 TTGCAGTTAGTTATGATGTCCCGTTATCAATCTAAGGAATTTGCAGAATTA  
 ACTTGGTCCATTATGGCAGCAAATGGTATTAAAGCCTTATATTTA

MSHMNYKEIYQEWLEND<sup>s</sup>SLGKDIKSDLEAIKGDSEIQDRFYKTLEFGTAGLR  
 KGLGAGTNRMTY<sup>s</sup>VMVGKAAQALANRLLIMALKLLHVELQLVMMSRYQSKE  
 FAELTWSIMAANGIKALYL

#### Sequence description:

A] Length: 366 bp - 122 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence preceded the 'ATG' start codon. No obvious potential leader peptide sequence.

ID-179

Clone 2-5d (ID-112d)

ATGCAACCTGTAAAAGTCGATGAACCTTCTGTTGAAGAAACCATTACTATT  
 TTGAAAGGTATCCAAAAAATACGAAGATTATCATCACGTAAAAATATAA  
 TAATGATGCCATAGAAGCAGCTGCAGTACTATCTAATCGTTATATCCAAGA  
 CCGCTTTTACCTGATAAAGCAATAGACTTATTAGATGAAGCTGGTTCTAA  
 AATGAACCTAACACTAAATTTTGGTTGATCCAAAAGAAATTGATCAACGTCT  
 CATTGAAGCAGAAAATTTAAAAGCGCAAGCGACTCGTGAAGAAGATTACG  
 AACGTGCAGCTTACTTCCGTGACCAGATTGCAAAATATAAAGAAATGCAG  
 CAACAAAAGGTTCGACGATCAAGATACACCTATTATTACCGAAAAACAAT  
 TGAGCACATCATTGAAGAAAAAACGAATATCCCTGTTGGTGATTAAAAAG  
 AAAAAGAACAACTCAATTAATTAATCTCGCAGATGACTTGAAACAGCAT  
 GTGATCGGCCAGGATGACGCTGTCATTAAGATTGCAAAAGCTATTTCGTCGT  
 AATCGAGTTGGTCTTGGTAGCCAAACCGTCCTATTGGTTCCCTTTTTATTG  
 TAGGACCAACCGGTGTTGGTAAACTGAACTTTCTAAACAAGCTAGCAATTG  
 AGCTCTTTGGTTCAGCTGATAGTATGATTCGTTTTGATATGTCAGAGTACAT  
 GGAAAAGCATGCTGTTGCTAAATTAGTCGGAGCGCCTCCAGGATACGTGG  
 GATACGAGGAAGCTGGACAATACTGAAAAGGTTTCGTCGAAATCCTTAC  
 TCGCTCATCCTTCTAGATGAAATTGAAAAAGCTCATCCCGATGTCATGCAT

FIG. 1<sub>CONT'D</sub>

FIG. 1 CONT'D

FIG. 1 CONT'D

84 / 110

Sequence description:

- A) Length: 498 bp - 165 aa (Partial gene sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-117 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-117 gene sequence. N- and C-termini have yet to be determined

ID-182

Clone 3-8b (ID-120b)

ATGTACCATATTGAATTAAAAAAGGAAGCTTTACTACCAAGAGAACGCCT  
AGTTGATTTAGGCGCAGATAGATTGAGTAATCAGGAGTTATTAGCCATTCT  
CTTACGTACAGGTATTAAGAAAAACCTGTTCTTGAAATTTCAACGCAAAT  
TTTAGAAAACATAAGCAGTTTAGCAGATTTTGGTCAATTATCCTTACAGGA  
GTTGCAATCCATTAAAGGAATCGGTCAGGTTAAATCCGTCGAAATAAAAG  
CTATGCTAGAACTAGCAAAACGGATTACAAAGCTGAATATGATCGTAAA  
GAGCAAATTTTAAGTAGTGAACAATTAGCGAGGAAAATGATGCTCGAATT  
AGGGGATAAAAAACAAGAACATTTAGTAGCTATTTATATGGATACACAAA  
ATCGTATTATCGAACAGAGAACTATTTTATTGGTACTGTACGTCGTTTACG  
TAGCAGAGCCAAGAGAAATTCTACATTATGCTTGTAACAAATCCGCAACT  
TCTTTGATTATTATACATAATCATCCCTCAGGTTCTCCAAATCCCAGTGAAA  
GTGATTTAAGTTTCACTAAAAAAATAAACGATCATGTGATCATCTGGGAA  
TTGTCTGCCTAGATCACATCATCGTTGGAAAAAATAAATATTATAGTTTTT  
GAGAAGAAGCAGATATTTTATAA

MYHIELKKEALLPRERLVDLGADRLSNQELLAILLRTGIKEKPVLEISTQILENI  
SSLADFGQLSLQELQSIKGIGQVKSVEIKAMLELAKRIHKAHEYDRKEQILSSEQ  
LARKMMLELGDKKQEHVLAIFYMDTQNRHIEQRTIFIGTVRRSVAEPREILHYAC  
KNMATSLIIHNHPSGSPNPSESDLSFTKKIKRSCDHLGIVCLDHIVGKNKYYSF  
REEADIL\*

Sequence description:

- A) Length: 681 bp - 227 aa (full-length gene)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-120 gene which was identified by LEEP,

FIG. 1 CONT'D

85 / 110

during cloning and sequence analysis of the full-length ID-120 gene sequence.  
ATG start codon is preceded by an typical  
Shine-Dalgarno sequence. No obvious leader  
peptide sequence

ID-183

Clone 3-11b (ID-121b)

TGGTTAAAAGTAGTGATAGCTTGTATTCCATCTATTTTAATTGCTTTACCAT  
TTGATAATTGGTTTGAAGCTCATTTTAATTTTCATGATTCCGATTGCAATAGC  
CCTAATCTTTTATGGTTTTGTCTTCATATGGGTTGAAAAACGTAATGCACAC  
CTCAAACACAGGTAACCGAATTGGCAAGTATGTCTTACAAGACAGCTTTC  
TTGATTGGATGTTCCAGGTTCTCAGTATTGTTCCGGGAACAGTCGTTCTG  
GAGCTACTATTTTAGGAGCAATTATTATTGGAAGTATGTCGTTCCGGTCGCTG  
CTGACTTTACTTTCTTCCTTGCCATCCCAACTATGTTTGGTTATAGTGGACT  
TAAGGCGGTTAAATATTTTTTAGATGGTAACGTCTTGAGTTTAGACCAATC  
TTTAATACTTTTAGTAGCAAGTCTGACAGCTTTCGTAGTTAGTTTATATGTT  
ATTCGTTTCTTGACAGACTATGTCAAACGACACGATTTACCATCTTTGGT  
AAGTATCGTATAGTCTTAGGAAGTTTACTCATCCTCTACTGGTTAGTTGTTT  
ATTTATTCTAA

WLKVVIACIPSILIALPFDNWFEAHFNFMIPAIALIFYGFVFIWVEKRNAHLKP  
QVTELASMSYKTAFLIGCFQVLSIVPGTSRSGATILGAIIGTSRSVAADFTFFLA  
IPTMFGYSLKAVKYFLDGNVLSLDQSLILLVASLTAFFVVSLEYVIRFLTDYVKR  
HDFTIFGKYRIVLGSLLILYWLVVHLF\*

Sequence description:

- A) Length: 579 bp - 193 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-68 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-68 gene sequence described in WO 00/06736. N-terminus has yet to be determined.

ID-184

Clone 3-11c (ID-121c)

FIG. 1<sub>CONT'D</sub>

86 / 110

ATGGAAATGAAACAAATCAGTGAAACAACACTGAAAATTACAATTAGTAT  
GGAAGATTTAGAAGATCGTGGTATGGAGCTGAAAGATTCCTAATCCCTCA  
GGAGAAGACTGAGGAATTTTTCTATTCTGTCATGGATGAATTAGACTTGCC  
AGAAAACCTTTAAAAATAGTGGTATGTTAAGTTTTTCGAGTAACACCTAAAA  
AAGATCGCATTGATGTTTTTGTACAAAGTCTGAATTAAGTAAAGATTAA  
ATTTAGAAGAATTAGCAGATTTGGGTGACATTTCAAAAATGTCTCCAGAAG  
ACTTTTTTAAAACCTTGGAACAATCGATGTTGGAAAAAGGGGATACGGAT  
GCCCATGCCAAATTAGCAGAAATTGAAAATATGATGGATAAAGCAACTCA  
AGAAGTAGTTGAGGAAAATGTTTCTGAAGAACAACCTGAAAAGGAAGTAG  
AAACGATTGGATATGTTCACTATGTCTTTGATTTTGATAATATTGAAGCTGT  
AGTTCGATTTTCAAAACGATTGATTTTCCAATAGAAAGCTT

MEMKQISETTLKITISMEDLEDGRGMELKDFLIPQEKTEEFFYSVMDELDLPENF  
KNSGMLSFRVTPKKDRIDVFVTKSELSKDLNLEELADLGDISKMSPEDFFKTLE  
QSMLEKGD TDAHAKLAEIENMMDKATQEVVEENVSEEQPEKEVETIGYVHY  
VFDFDNIEAVVRFSQTIDFPIEA

Sequence description:

- A) Length: 547 bp - 182 aa (Partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-68 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-68 gene sequence. ATG start codon is preceded by an typical Shine-Dalgarno sequence. No obvious potential leader peptide sequence

ID-185

Clone 3-16b (ID-122b)

GGAAACCAACGGCCAGTACAATCGTCAAGGGTAGATTATCCTAAACGTAG  
TCGTGCCAAGATTGTAGAAGTTTATTTTAGACAAGCTTCTACTACTGATTA  
TTCTGGTGTTTACAAAGGTTACTATATTGACTTTGAAGCCAAAGAAACCCG  
GCAGAAAACCTGCTATGCCTATGAAAAATTTTCATGCTCACCAAATAGAGC  
ACATGGCAAATGTATTACAGCAAAAAGGGATTGCTTTGTCTTGCTTCATT

FIG. 1 CONT'D

87 / 110

TTTCCACACTTAAGGAAACCTATCTACTCCCTGCTAATGAGTTAATTTTCATT  
TTATCAGATTGATAAAGGCAATAAATCAATGCCTATTGATTATATCAGAAA  
AAATGGATTTTTCGTAAAGGAGAGTGCCTTTCCTCAAGTCCCTTACTTAGA  
TATTATTGAAGAAAAATTATTAGGCGGTGATTACAATTAA

GNQRPVQSSRVDYPKRSRAKIVEVYFRQASTTDYSGVYKGYIDFEAKETRQ  
KTAMPMKNFHAHQIEHMANVLQKGICFVLLHFSTLKETYLLPANELISFYQI  
DKGNKSMPIDYIRKNGFFVKESAFPQVPYLDIIEKLLGGDYN\*

Sequence description:

- A) Length: 447 bp - 149 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-122 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-122 gene sequence. N-terminus has yet to be determined

ID-186

Clone 3-17b (ID-123b)

GGATCCTAAAAACGCTAAGGTTTATCAAAAAAATGCTGATCAATTTAGTG  
ACAAGGCAATGGCTATTGCAGAGAAGTATAAGCCAAAATTTAAAGCTGCA  
AAGTCTAAATACTTTGTGACTTCACATACAGCATTCTCATACTTAGCTAAG  
CGATACGGATTGACTCAGTTAGGTATTGCAGGTGTCTCAACCGAGCAAGA  
ACCTAGTGCTAAAAAATTAGCCGAAATTCAGGAGTTTGTGAAAACATATA  
AGGTTAAGACTATTTTGTGTTGAAGAAGGAGTCTCACCTAAATTAGCTCAAG  
CAGTAGCTTCAGCTACTCGAGTTAAAATTGCAAGTTTAAGTCCTTTAGAAG  
CAGTTCCCAAAAACAATAAAGATTACTTAGAAAATTTGGAAACTAATCTTA  
AGGTACTTGTCAAATCGTTAAATCAATAG

DPKNAKVYQKNADQFSDKAMAIAEKYKPKFKAASKYFVTSHTAFSYLAKR  
YGLTQLGIAGVSTEQEPSAKKLAEIQEFVKTYKVKTIFVEEGVSPKLAQAVAS  
ATRVKIASLSPLEAVPKNNKDYLENLETNLKVLVKS LNQ\*

Sequence description:

FIG. 1<sub>CONT'D</sub>

88 / 110

A] Length: 433 bp - 144 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-123 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-123 gene sequence. N-terminus has yet to be determined

ID-187

Clone 3-46/47 (ID-130b)

ATGAAAAAAGTCATCGATTTAAAAAACTACAAAAAGCATACGCCTCAGA  
AACTGTTTTAAATAATATTAATTTGGAGGTGTTTAAAGGAGAAATAATTGG  
ATTAATAGGACCCTCTGGAGCAGGGAAATCTACCTTGATTAAAACTATGCT  
TGGCATGGAAAAAGCAGATAAGGGAACAGCTCTTGTTCTTGATACTCAAA  
TGCCAGATCGTAATATTTTAAATCAAATTGGCTATATGGCTCAATCTGATG  
CCTTACACGAGTCTTTAACTGGCTTAGAAAAATTATTATTCTTTGGAAAAA  
TGAAAGGTATTCAAAAAACTGAATTAACACAGCAGATAACTCATATTTCT  
AAAGTAGTAGATCTAGAAAACCAACTTGATAAATTTGTCTCAGGTTACTCA  
GAAGGTATGAAAAGACGGCTTTCTCTAGCCATCGCCCTACTTGGAACCCC  
ACAGTTTTAATCCTAGATGAACCTACCGTTGGAATTGATCCATCCTTGAGG  
AGAAAAATCTGGCAAGAGCTAATTAATATTAAGGATGAAGGACGTTCTAT  
CTTTATTACAACCCACGTTATGGATGAAGCAGAATTAACAAGTAAGGTTGC  
ACTACTATTACGTGGAAACATTATTGCCTTTGATACTCCATTACATTTAAA  
AAAACAATTTAATGTGAGTACTATTGAGGAAGTTTTCTTAAAAGCTGAAGG  
AGAATAA

MKKVIDLKKLQKAYASETVLNNINLEVFKGEIIGLIGPSGAGKSTLIKTMLGME  
KADKGTALVLDTPMPDRNINLQIGYMAQSDALHESLTGLENLLFFGKMKGIQ  
KTELKQQITHISKVVDLENQLDKFVSGYSEGMKRRLSLAIALLGNTPTVLILDEP  
TVGIDPSLRRKIWQELINIKDEGRSIFITTHVMDEAELTSKVALLLRGNIIAFDTP  
LHLKKQFNVSTIEEVFLKAEGE\*

Sequence description:

A] Length: 717 bp - 239 aa (Possible full-length sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-130 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-130 gene sequence. ATG start codon is preceded by a possible

FIG. 1<sub>CONT'D</sub>

89 / 110

Shine-Dalgarno. No obvious potential leader  
peptide sequence

ID-188

Clone 3-83b (ID-144b)

ATGGTACAAATGATACATGATATGATTAAACAATTGAGCATTGCTGAG  
ACACAAGCTGATTTTCCAGTGTATGATATTTAGGGGAAGTCCATACTTAT  
GGACAACCTAAAGTAGACTCTGACTCTCTAGCTGCTCATATTGATAGCCTA  
GGCCTTGTTGAAAAATCACCTGTCTTAGTATTCGGTGGTCAAGAATATGAA  
ATGTTGGCGACATTTGTTGCTTTAACAAAGTCAGGGCATGCTTATATACCG  
GTTGACCAACACTCTGCTTTGGATAGATAACAGGCTATTATGACAGTTGCT  
CAACCAAGCCTTATCATTTC AATTGGTGAATTTCTCTTGAAGTTGATAAT  
GTCCCAATCCTAGACGTTTCTCAAGTTTCAGCTATTTTTGAAGAAAAGACT  
CCTTATGAGGTAACACATTCTGTAAAGGTGATGATAATTACTATATTATT  
TTCATTTCAGGGACTACTGGTTTACCAAAAAGGTGTGCAAATTTACATGAC  
AATTTATTGAGCTTTACAAATTGGATGATTTCTGATGATGAGTTTTCAGTTC  
CTGAAAGACCGCAAATGTTGGCTCAACCC

MVQMIHDMIKTIEHFAETQADFPVYDILGEVHTYGQLKVDSDSLAAHIDSLGL  
VEKSPVLVFGGQEYEMLATFVALTKSGHAYIPVDQHSALDRIQAIMTVAQPSL  
IISIGEFPLEVDNVPILDVSQVSAIFEKTPYEVTHSVKGDDNYIIFTSGTTGLP  
KGVQISHDNLLSFTNWMISDDEFSVPERPQMLAQP

Sequence description:

- A) Length: 592 bp - 197 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-144 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-144 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. No obvious leader peptide sequence  
This orf is not in frame with nuc

ID-189

FIG. 1<sub>CONT'D</sub>

90 / 110

Clone 3-86b (ID-145b)

ATGGAAAATCATCGTTATGAAGATGAAGGTAAATTCCAGCGTAAGATGAC  
CAGTCGTCATCTCTTTATGTTATCGCTAGGTGGTGTATCGGGACTGGGCTT  
TTCTTGAGTTCAGGTTATACCATTGCACAGGCTGGTCCGCTTGGAGCTGTG  
CTGTCTTATTTGATTGGTGCCGTTGTGGTTTATTTGGTCATGCTATCACTTG  
GGGAATTGGCGGTTGCCATGCCGGTGACGGGGTCATTCCACACTTATGCCA  
CTAAGTTTATCAGTCCTGGAACAGGTTTTACTGTTGCTTGGCTATATTGGAT  
TTGTTGGACGGTCGCCTTGGGGACTGAATTTTTAGGTGCTGCCATGCTGAT  
GCAGCGCTGGTTCCCAAATGTGCCGGCTTGGGCATTTGCTTCTTTTTGCC  
CTTGATTTTTGGTTTAAATGCTCTTAGCGTACGCTTTTTTGCAGAAAGCAG  
AGTCTTCTCTCAAGTATTAAGGTTATTGCTATCATTATCTTTATTATCTTG  
GGCTTAGGTGCTATGTTTGGTCTAGTTTCCTTTGAAGGTCAGCACAAAGGCT  
ATTCTCTTCACTCATCTGACTGCCAATGGTGCCTTCCCAAATGGTATCGTTG  
CAGTTGTCTCAGTCATGTTGGCTGTTAACTATGCCTTCTCTGGTACTGAGTT  
AATTGGTATTGCCGGCTGGTGAAACGGATAATCCCAAAGAAGCTGTACCAA  
GGGCTATTAAAACGACAATCGGTGCGTTGGTTGTTTTCTTTGTACTGACAA  
TTGTTGTCCTAGCTTCGCTATTGCCAATGAAAGAGGCAGGCGTATCCACAG  
CACCATTTCGTTGATGTCTTTGACAAGATGGGAATCCCTTTTACGGCGGATA  
TCATGAACTTCGTTATCTTGACAGCCATCCTGTCTGCTGGTAACTCAGGTCT  
CTACGCATCAAGCCGTATGCTCTGGTCCCTTGCCAATGAAGGTATGTTGTC  
AAAATCTGTTGTGAAAATCAATAAACACGGGTGTCCTCAATGCGTGCTCTTCT  
CTTGTCATGGCAGGAGCAGTGCTGTGCTCTTTTCAAGTATTTACGCTGC  
AGACACAGTTTATCTAGCCTTGGTTTCAATCGCGGGCTTTGCTGTTGTTGTC  
GTATGGCTAGCCATTCCAGTCGCACAAATCAATTTCCGCAAGGAATTC

MENHRYEDEGKFQRKMTSRHLFMLS LGGVIGTGLFLSSGYTIAQAGPLGAVL  
SYLIGAVVVYLVMLSLGELAVAMPVTSFHTYATKFISPGTGFTVAWLYWIC  
WTVALGTEFLGAAMLQMQRWFPNVPWAFASFFALVIFGLNALSVRFFAEAES  
FFSSIKVIAIIFILGLGAMFGLVSFEGQHKAILFTHLTANGAFPNGIVAVVSVM  
LAVNYAFSGTELIGIAAGETDNPKEAVPRAIKTTIGRLVVFVLTIVVLASLLPM  
KEAGVSTAPFVDVFDKMGIPFTADIMNFVILTAILSAGNSGLYASSRMLWSLA  
NEGMLSKSVVKINKHGVPMRALLLSMAGAVLSLFSSIYAADTVYLALVSIAGF  
AVVVVWLAIPVAQINFRKEF

Sequence description:

A) Length: 1126 bp - 393 aa (partial gene)

FIG. 1<sub>CONT'D</sub>

91 / 110

sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-145 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-145 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a possible leader peptide sequence.

ID-190

Clone 3-94b

TCAGAAAATGCAGAGGCAGCAACGGTTGCCACAACTTGGTTACCAAAGG  
AGCTAATGTCATTATCGGACCAGCAACATCGGGTGCAGCTGCATCTTCAAC  
TCCAAAAGTAAATGCAGCAGCAGTTCCAATGATTGCACCTGCTGCGACAC  
AAGACAATTTAGTCTATGGTTCTGATGGAAAAACCTTAAATCAGTATTTCT  
TCCGAGCTACTTTTGTCTGATAATTATCAAGGAAAGCTATTGTCTCAGTATG  
CTACAGACAACCTTAAAGCTAAAAAAGTTGTTCTATTTTATGATAATTCAT  
CAGATTACTCAAAGGGGGTAGCAAAATCATTTAAGGAAAGTTATAGTGGA  
AAAATTGTTGATAGTATGACATTCTCCGCTGGTGATACTGATTTCCAAGCG  
TCATTGACTAAGTTGAAAGGGGAAAGAATATGATGCTATTGTGATGCCAGG  
TTACTATACCGAGACAGGATTAATAGTTAAGCAAGCGCGTGATTTAGGTAT  
CTCTAAACCGGTTCTTGGGCCTGATGGTTTTGATAGTCCGAAATTTGTGCA  
ATCGGCAACACCTGTAGGAGCTTCAAACGTTTATTATTTGACAGGTTTCAC  
TACACAAGGATCAACCAAAGCTAAAGCT

SENAEAATVATNLVTKGANVIIGPATSGAAASSTPKVNAAA VPMIAPAATQD  
NLVYGSDGKTLNQYFFRATFVDNYQGKLLSQYATDNLKAKKVVLFDNSSD  
YSKGVAKSFKESYSGKIVDSMTFSAGD TDFQASLTKLKGKEYDAIVMPGYT  
ETGLIVKQARDLGISKPVLPDGF DSPKFVQSATPVGASN VYYLTGFTTQGST  
KAKA

Sequence description

A] Length: 637 bp - 231 aa (partial sequence)

FIG. 1 CONT'D

92 / 110

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-149 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-149 gene sequence. N- and C-termini have yet to be determined

ID-191

Clone 2-c94b (ID-153b)

TTGGGACTTAAAGACCATGCTTTAGTCTATCCATTTTCATTATCTGGGGGG  
CAAAAGCAACGTGTCGCACTAGCTCGTGCGATGATGATTGATCCACAGATT  
ATTGGTTATGATGAGCCAACTAGCGCTCTTGATCCAGAGTTGCGTCAAGAA  
GTAGAAAACTAATTTTACAAAATAGAGAAACAGGTATGACACAAATTGT  
AGTAACACATGATCTTCAATTTGCTGAAAGTATATCTGATACGATTCTCAA  
AATTAATCCTAAGTAG

MGLKDHALVYPFSLSGGQKQRVALARAMMIDPQIIGYDEPTSALDPELRQEV  
EKLILQNRETGMTQIVVTHDLQFAESISDTILKINPK\*

Sequence description

A] Length: 270 bp - 90 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-153 gene which was identified by LEEP, during cloning and sequence analysis of the ID-153 gene sequence.

N-terminus has yet to be determined

ID-192

Clone 2-c1b (ID-155b)

ATGACTAATATCTCAGATGTTCCAAAAGCTATTAGAACACAGGCACAGTAT  
GTTCTCTTGGGAATGAGAGTTATGGATCAGTCGGTATTACCGAAAACATAT  
AATTCAAAAGAACCTTATTTGAAACCAGATATGATTTATATTCATGATAGA

FIG. 1 CONT'D

93 / 110

AGACAAGAGACAATGCTTAAAATCACTCAAGAAATAGAAATGGAGCATTG  
A

MTNISDVPKAIRTQAQYVLLGMRVMDQSVLPKTYNSKEPYLKPDMIYIHDRR  
QETMLKITQEIEMEH\*

#### Sequence description

- A) Length: 204 bp - 68 aa (partial sequence)
- B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-155 gene which was identified by LEEP, during cloning and sequence analysis of the ID-155 gene sequence.
- ATG start codon is preceded by a potential typical Shine-Dalgarno sequence.
- Has a typical leader peptide. N-terminus has yet to be verified

#### ID-193

Clone 2-54altb (ID-172b)

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCTTGGGGAATATAAATT  
TGGATTTTCATGACGATGTAAAGCCAATTTATTCTACGGGAAAAGGTCTAAA  
TGAGGCTGTTATTCGTGAGTTATCTGCAGCTAAGGGTGAACCTGAGTGGAT  
GTTGGACTTTCGTCTAAAATCCTTGGAAACGTTTAATAAAAATGCCGATGCA  
GACCTGGGGAGCAGATTTATCAGATATTGATTTTGATGATATTATTTATTA  
TCAAAAAGCATCTGATAAACCTGCGCGTGATTGGGATGATGTTCCAGAAA  
AAATCAAAGAAACTTTTGAAAGAATTGGGATTCCAGAAGCTGAAAGAGCC  
TATCTTGCAGGAGCATCAGCACAATATGAATCAGAAGTAGTTTATCACAAT  
ATGAAAGAAGAATATGATAAGCTGGGTATTGTTTTTACGGATACTGACTCT  
GCACTTAAAGAGTACCCAGAGCTATTCAAAAAATATTTTGCTAAACTTGTC  
CCTCCAACAGATAATAAATTAGCTGCTCTGAACTCTGCTGTATGGTCAGGT  
GGAACATTTATTTATGTTCTAAAGGTGTTAAGGTGGATATTCCACTTCAA  
ACTTACTTCCGTATTAATAATGAAAATACTGGACAATTTGAACGTACTCTC  
ATTATTGTTGATGAGGGAGCAAGTGTTCACTATGTTGAAGGTTGTACCGCC  
CCAACCTTATCTTCAAATAGTTTACATGCAGCTATAGTTGAAATTTTGCAC  
TTGATGGAGCTTATATGCGCTATACGACTATTCAAAATTGGTCCGATAATG  
TCTATAATTTAGTGACAAAACGTGCTACCGCTAAAAAAGATGCAACAGTT  
GAGTGGATAGATGGAAATCTAGGAGCTAAACAACAATGAAATACCCATC

FIG. 1<sub>CONT'D</sub>

GGTTTACCTTGATGGTGAAGGAGCACGTGGCACGATGTTGTCTATTGCTTT  
 TGCAAACAAAGGACAACACCAAGATACGGGTGCAAAGATGATTCATAATG  
 CCCCCATACTAGTTCATCCATTGTCTCTAAATCAATTGCTAAGGGTGGGG  
 GAAAAGTTGATTATCGAGGTCAAGTGACATTTAATAAAGATTCCAAAAA  
 TCAGTGTACATATAGAATGTGACACCATATTGATGGATGATATTTCAAAA  
 TCAGATACCATAACCGTTTAATGAGATTCATAATTACAGGTTGCTTTAGAG  
 CATGAAGCAAAGGTGTCTAAGATTTCTGAAGAGCAACTGTACTACTTGATG  
 AGTCGAGGTTTATCTGAAGCTGAAGCAACAGAAATGATTGTTATGGGGTTT  
 GTTGAGCCCTTTACGAAAGAATTACCAATGGAATATGCGGTAGAGTTAAA  
 TCGTTTAATTTCTATGAAATGGAAGGTTGAGTTGGTTAA

MHACRSTLEDLGEYKFGFHDDVKPIYSTGKGLNEAVIRELSAAKGEPEWMLD  
 FRLKSLETFNKMPMQTWGADLSIDFDDIYYQKASDKPARDWDDVPEKIKE  
 TFERIGIPEAERAYLAGASAQYSEVVYHNMKEEYDKLGVFTDSDSALKEYP  
 ELFKKYFAKLVPPTDNKLAALNSAVWSGGTFIYVPKGVKVDIPLQTYFRJNE  
 NTGQFERTLIIVDEGASVHYVEGCTAPTYSSNSLHAAIVEIFALDGAYMRYTTI  
 QNWSDNVYNLVTKRATAKKDATVEWIDGNLGAKTMMKYPSVYLDGEGARG  
 TMLSIAFANKGQHQDTGAKMIHNAPHTSSSIVSKSIKGGGKVDYRGQVTFN  
 KDSKKSVSHECDTILMDDISKSDTIPFNEIHNSQVALEHEAKVSKISEEQLYYL  
 MSRGLSEAEATEMIVMGFVEPFTKELPMEYAVELNRLISYEMEGSVG\*

Sequence description:

- A] Length: 1411 bp - 469 aa (Possible full-length gene)
- B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-72 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-72 gene sequence. No obvious Shine Dalgarno sequence upstream of TTG start codon (insufficient sequence data). N terminus needs verification.

ID-194

Clone 3-1b (ID-81b)

ATGATAGAATTCTTTTCTAATATCAGAACAGAGATTCCGCAGATGCCTTTA  
 CTTATCCATAGTTTGATTTTATCTGTCTTACCTTTTCTGATGTGGCTGACTTT  
 GGTTAATAGAGATAAGCCTTTGTATAAACTATTTGGAGTATCCTTTTAGG  
 ACTTCAGTTAATTACGATTTATACTTGGTTTTTCTGGGCAAATTGCCTTTA

FIG. 1 CONT'D

95 / 110

TCTGAAAGTCTTCCCCTTTACCATTGTCTGAATAGGCATGTTTGTCGGTCTCT  
TA

MIEFFSNIRTEIPQMPLLIHSLILSVLPFLMWLTLVNRDKPLYKTIWSILLGLQLI  
TIYTWFFWAKLPLSESLPLYHCRIGMFVGLL

#### Sequence description

A) Length: 261 bp - 87 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-81 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-81 gene sequence. Sequence Characteristics: Possesses a potential leader peptide sequence. Orf is preceded by a potential Shine-Dalgarno sequence.

ID-195

Clone RS-55b

AAGCTTGTGCAAAGTATTAAAGAGATAGGATTAGCTAATGCGCATTATTATTA  
GCTGTTGCTCCGACAGGGTCAATCAGTTATCTTCTTCTTGTAAGTCCGAGCC  
TTCAACCGGTTGTATCACCTGTCTGAAGTACGCAAGGAAGGAGCACTGGGG  
AGGGTTTATGTAGCTGCTTATAAGATTGATGCAGATAATTATGTCTACTAC  
AAAAAAGGAGCTTATGAAGTGGGATCTGAGGCGATTATCAATATTGCAGC  
TGCCGCTCAAAAACACATTGATCAAGCTATTTTCGTTAACGCTTTTCATGAC  
AGATCAAGCAACTACGCGAGATTTAAATAAAGCCTATATTCAAGCATTTA  
AACAAAAATGTGCCTCTATTTATTATGTACGAGTGAGACAGGACATCCTAG  
AAGGTAGCGAGAGTTATGATGATATGCTGGATGATTTCACTTCATCGGACT  
TAGAAGACTGTCAATCCTGCATGATTAA

>KLVQSIKEIGLANAHLAVAPTGSISYLSSCTPSLQPVVSPVEVRKEGALGRV  
YVAAYKIDADNYVYYKKGAYEVGSEAIINIAAAAQKHIDQAISLTLFMTDQAT  
TRDLNKAYIQAFKQKCSIYYVRVRQDILEGSESYDDMLDDFTSSDLEDCQSC  
MI\*

Sequence description:

FIG. 1<sub>CONT'D</sub>

96 / 110

A) Length 486 bp - 162 aa (Partial sequence)

B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-87 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-87 gene sequence. N-terminus to be determined.

ID-196

Clone RS-59(ID-90b)

GTGAGGACATATATTACAACTTGAATGGACATTCAATCACTAGTACAGC  
ACAAATAGCTCAAAACATGGTAACAGATATAGCAGTAAGCTTAGGTTTC  
GTGAGCTGGGAATACATTCTTATCCGATTGATACTGATTCTCCTGAGGAAA  
TGAGTAAGCGTTTAGATGGAATCTGTTCCGGACTTAGAAAAAATGATATTG  
TCATATTTTACAGACACCTACATGGAACACTACAACCTTTTATGATGAAAAATTAT  
TTCACAAATTAAAAATATTTGGTGTAAGATTGTTATTTTATACATGATGT  
TGTACCGCTAATGTTTATGATGGAAATTTTATTTGATGGATAGAACTATAGC  
TTATTATAATGAAGCAGATGTTTAATAGCCCCTAGTCAAGCAATGGTCGAT  
AAGCTT

MRTYITNLNGHSITSTAQIAQNMVTDIAVSLGFRELGIHSPIDTDSPEEMSKRL  
DGICSLRKNDIVIFQTPTWNTTTFDEKLFHKLKIFGVKIVIFIHDDVVPLMFDGN  
FYLMDRITAIYYNEADVLIAPSQAMVDKL

Sequence description:

A) Length: 414 bp - 138 aa(partial gene)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence. No obvious signal peptide, but a possible Shine Dalgarno sequence is present upstream of ATG start codon. C-terminus has yet to be determined.

ID-197

Clone RS-59c (ID-90c)

FIG. 1<sup>CONT'D</sup>

FIG. 1 CONT'D

FIG. 1 CONT'D

99 / 110

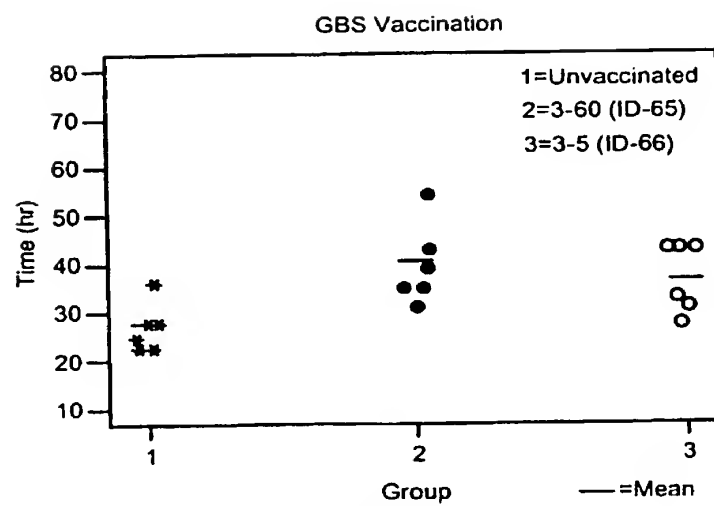


FIG. 2

nucS1

Bgl II Eco RV

5'-cgagatctgatatctcacaaacagataacggcgtaaataag -3'

nucS2

Bgl II Sma I

5'-gaagatcttcccccggatcacaaacagataacggcgtaaataag -3'

nucS3

Bgl II Eco RV

5'-cgagatctgatatccatcacaaacagataacggcgtaaataag -3'

nucR

Bam HI

5'-cgggatccttatggacctgaatcagcgttgtc -3'

NucSeq

5'-ggatgctttgtttcaggtgtatc -3'

pTREP<sub>F</sub>

5'-catgatatcggtacctaagctcatatcattgtccggcaatgggtgtgggctttttttgttttagcggataa  
caatttcacac -3'

pTREP<sub>R</sub>

5'-gcggatcccccggttaattaatgtttaaacactagtcgaagatctcgcaattctcctgtgtgaaatt  
gttatccgcta -3'

pUC<sub>F</sub>

5'-cgccagggttttccagtcacgac -3'

V<sub>R</sub>

5'-tcaggggggaggagcctatg -3'

V<sub>1</sub>

5'-tcgtatgttgtgtggaattgtg -3'

V<sub>2</sub>

5'-tccggctcgtatgttgtgtggaattg -3'

FIG. 3

pTREP-Nuc vectors allow cloning of genomic DNA into each frame with respect to the nuclease gene

(i)

pTREP1-nuc1 (EcoRV)	AAGTATCAGATCT-- <u>GATATC</u> --TCACAAACAGATAACGGCGTAAAT	Frame=+1
	.....▲.....	
pTREP1-nuc2 (Sma I)	AAGTATCAGATCTTCCCGGGA--TCACAAACAGATAACGGCGTAAAT	Frame=+2
	.....▲.....	
pTREP1-nuc3 (EcoRV)	AAGTATCAGATCT-- <u>GATATC</u> TCACAAACAGATAACGGCGTAAAT	Frame=+3
	.....▲.....	
Nuclease Gene	TCACAAACAGATAACGGCGTAAAT	

Cloning site is indicated by an arrow

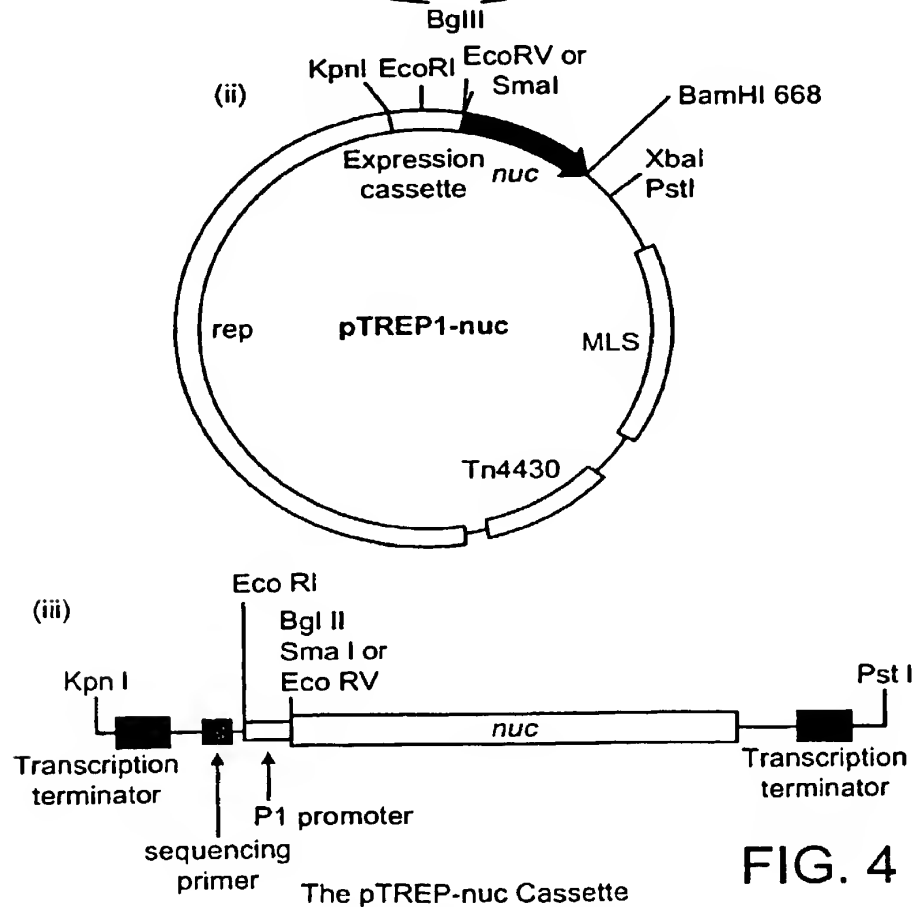
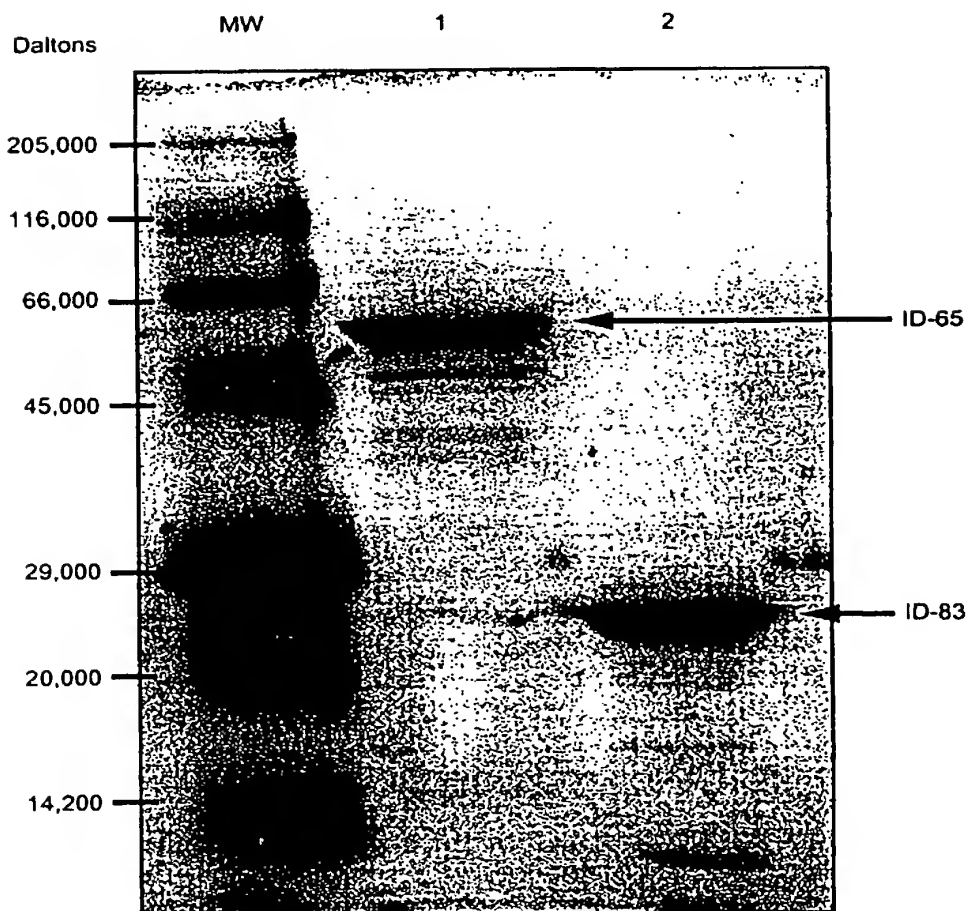


FIG. 4

102 / 110

# FIG. 5

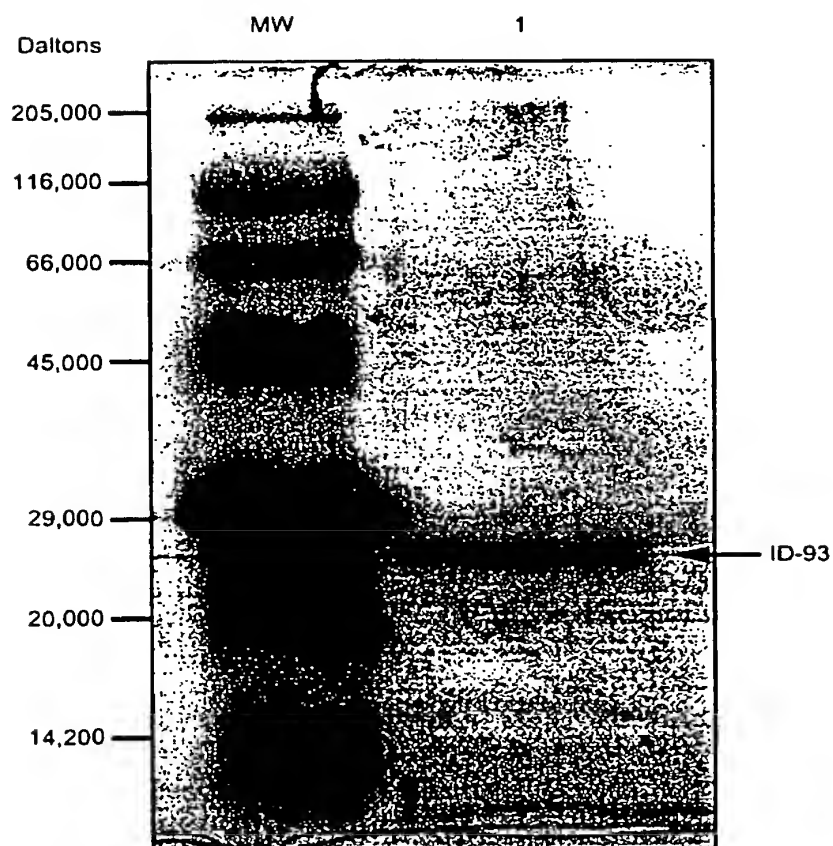
SDS-PAGE analysis of the purified ID-65 and ID-83 protein antigens



103 / 110

**FIG. 6**

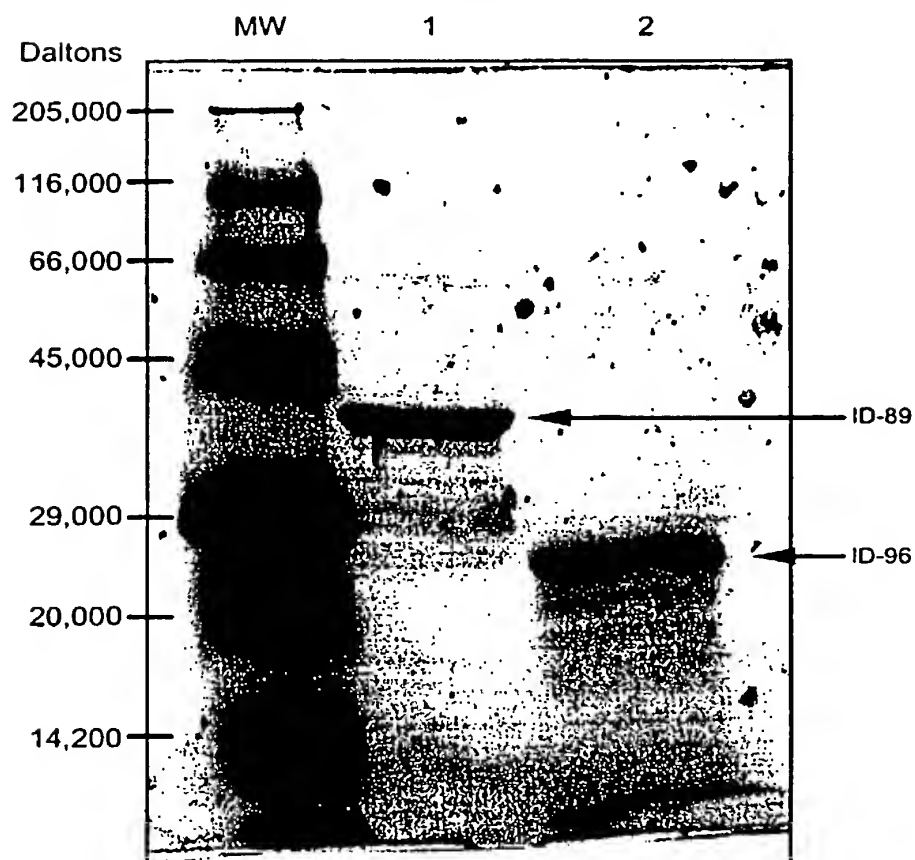
SDS-PAGE analysis of the purified ID-93 antigen



104 / 110

**FIG. 7**

SDS-PAGE analysis of the purified ID-89 and ID-96 protein antigens

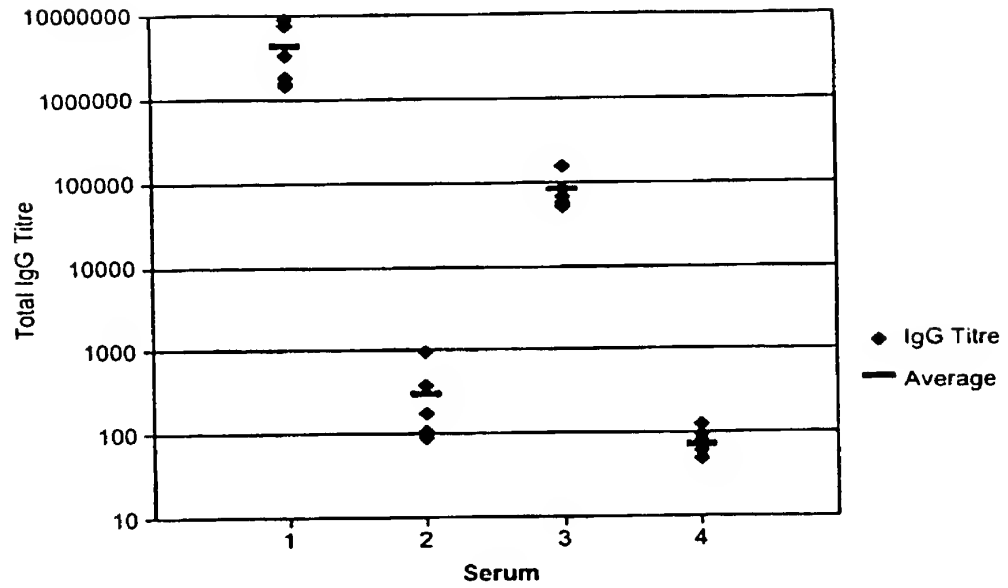


105 / 110

## FIG. 8

IgG Titres against the ID-65 and ID-83 proteins

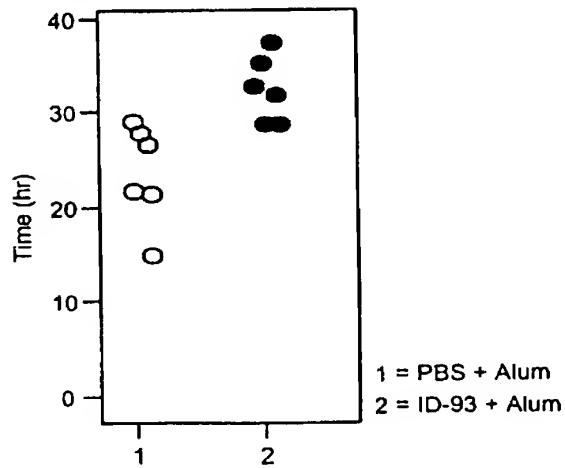
ID-65 and ID-83 Vaccinations -IgG Titres



## FIG. 9

Survival data

ID-93 Vaccination- GBS Challenge and Survival

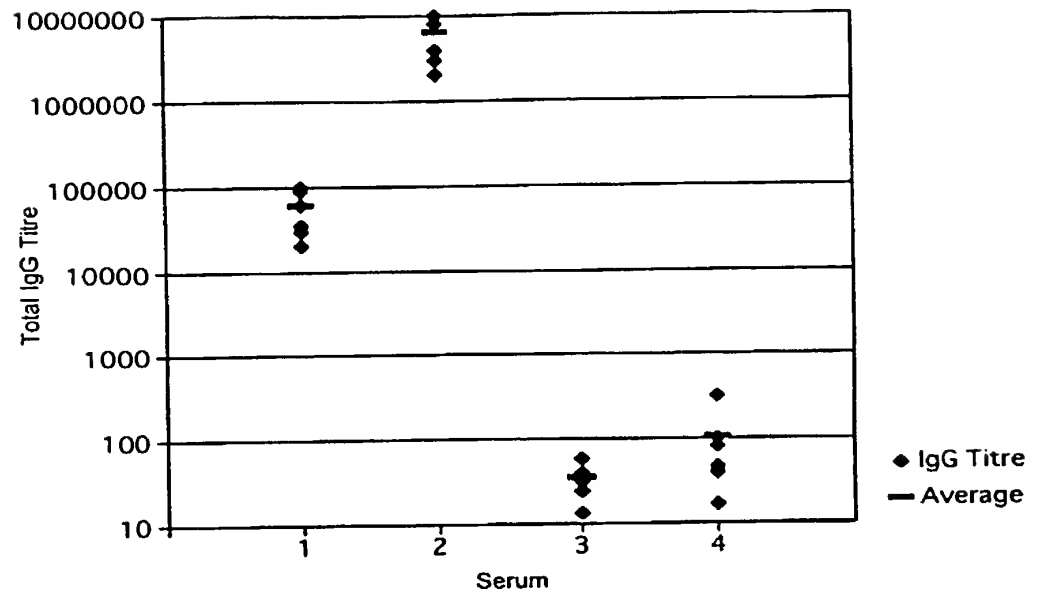


106 / 110

## FIG. 10

IgG Titres against the ID-93 protein

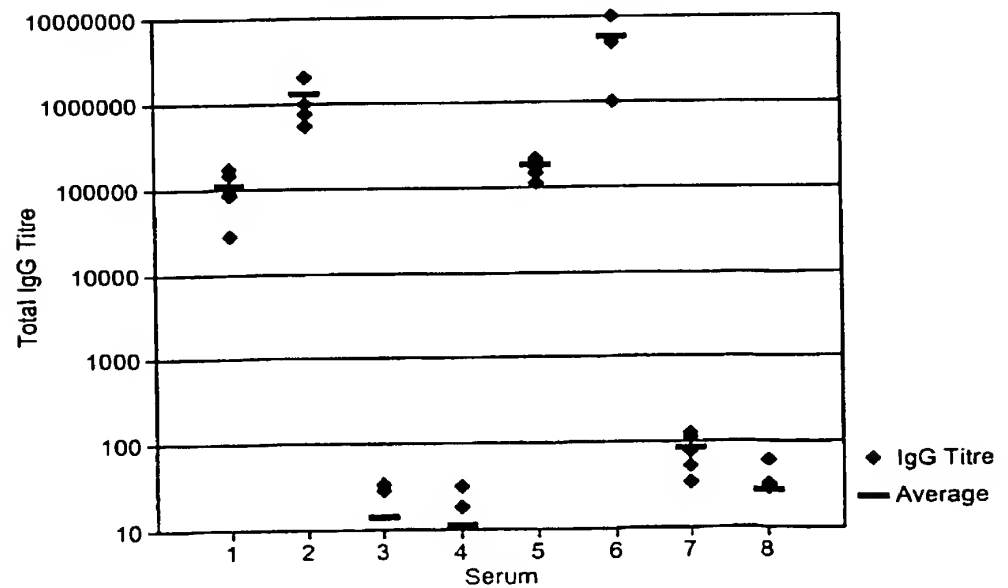
ID-93 Protein Vaccine -IgG Titres



## FIG. 11

IgG Titres against the ID-89 and ID-96 proteins

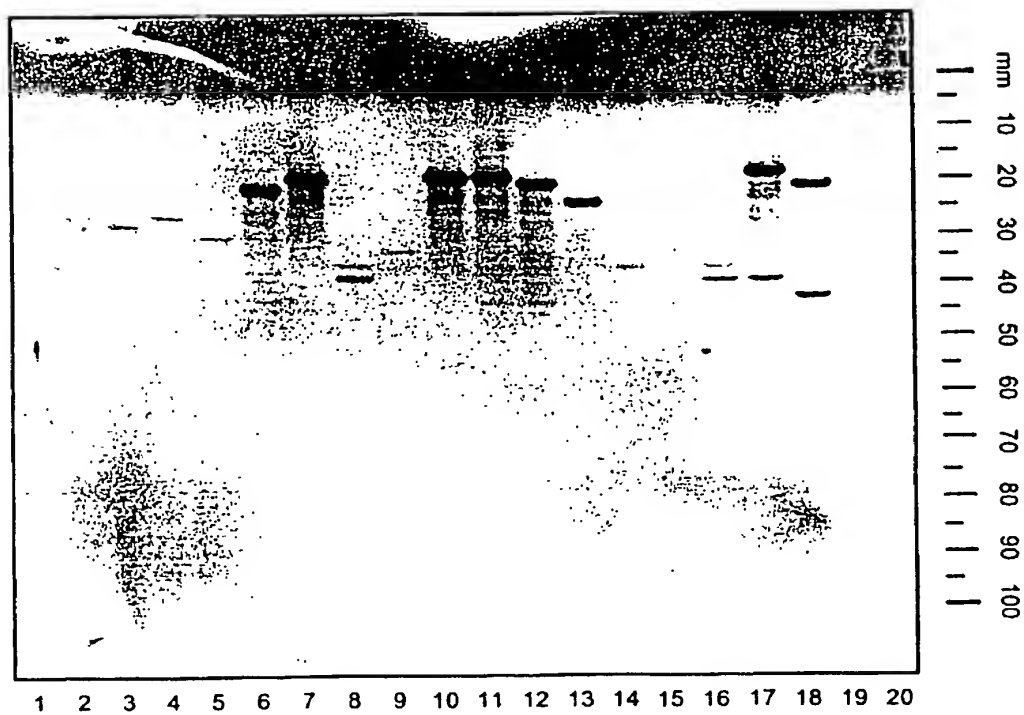
ID-89 and ID-96 Protein Vaccines -IgG Titres



107 / 110

# FIG. 12

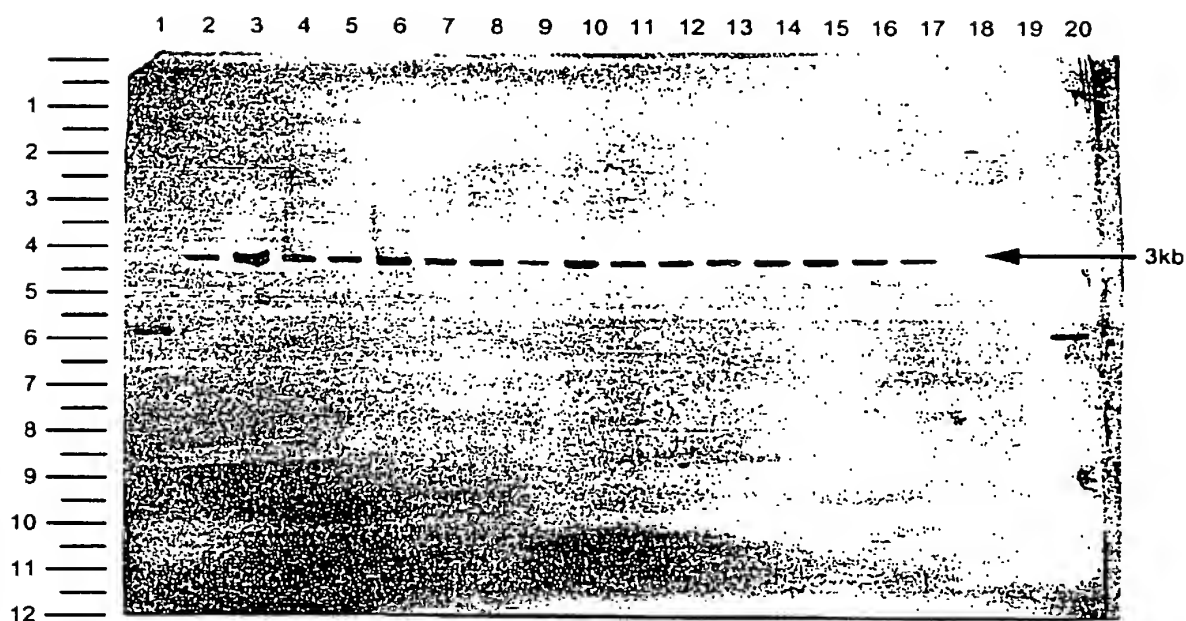
Southern blot analysis - *rib*



108 / 110

# FIG. 13

Southern blot analysis - ID-65



# FIG. 14

Southern blot analysis - ID-89

